To appear in *Cognitive Psychology*

Set-Shifting and Place-Keeping as Separable Control Processes

Richard P. Cooper, Catherine Byde, Roberto de Cecilio[1], Chelsea Fulks & Danila Morais

Department of Psychological Sciences,

Birkbeck, University of London

Abstract: We present three experiments using a sequential binary choice task that explore the relationship between two proposed cognitive control functions: set-shifting and place-keeping (i.e., keeping track of one's place within a sequential task). The task involves switching from one stimulus-response mapping to another across trials, according to a predefined sequence and in the face of occasional brief interruptions. Response-stimulus interval, interruption length and interrupting task were varied. The robust finding across all experiments was that varying response-stimulus interval led to standard effects attributable to set-shifting, while varying interruption length led to standard effects attributable to place-keeping, but in no cases did the factors interact. We interpret the results as supporting the view that set-shifting and place-keeping are achieved by separable control processes and illustrate this interpretation with a computational model of performance on the task.

Corresponding Author:

Richard P. Cooper
Department of Psychological Sciences
Birkbeck, University or London
Malet Street, London WC1E 7HX
United Kingdom

Phone: +44 20 7631 6211
Email: R.Cooper@bbk.ac.uk

Word Count: 17,966 (excluding title, references, author affiliations, acknowledgments, appendix, figures and figure legends)

Keywords: cognitive control; set-shifting; place-keeping; executive function; task interruption; procedural error

[1] Now at the University of Groningen.

Set-Shifting and Place-Keeping as Separable Control Processes

Richard P. Cooper, Catherine Byde, Roberto de Cecilio, Chelsea Fulks & Danila Morais

Department of Psychological Sciences,
Birkbeck, University of London

Abstract: We present three experiments using a sequential binary choice task that explore the relationship between two proposed cognitive control functions: set-shifting and place-keeping (i.e., keeping track of one's place within a sequential task). The task involves switching from one stimulus-response mapping to another across trials, according to a predefined sequence and in the face of occasional brief interruptions. Response-stimulus interval, interruption length and interrupting task were varied. The robust finding across all experiments was that varying response-stimulus interval led to standard effects attributable to set-shifting, while varying interruption length led to standard effects attributable to place-keeping, but in no cases did the factors interact. We interpret the results as supporting the view that set-shifting and place-keeping are achieved by separable control processes and illustrate this interpretation with a computational model of performance on the task.

1 Introduction

Almost 25 years ago Rogers and Monsell (1995) lamented the fact that, while it had "long been understood that cognitive processes require control processes to organize them", the mechanisms supporting that control were "to put it mildly – poorly understood" (p. 207). Much progress has been made in the intervening years (see, Logie, 2016, for a recent summary and discussion), yet the majority of studies have focused on one aspect of cognitive control or "executive function" (whether it be related to set-shifting, response inhibition, etc.), typically using relatively simple tasks. There is of course good reason for this, as initial progress would not have been possible without good levels of process purity (i.e., tasks whose performance primarily involves the operation of a single putative executive function) and high levels of experimental control over the many variables that affect performance on tasks that require elements of cognitive control.

Yet many tasks, particularly those outside of the laboratory, involve multiple putative cognitive control mechanisms working together if optimal (or even good) levels of performance are to be achieved. Indeed, even the Wisconsin Card Sorting Task (WCST: Milner, 1963), which is widely used in the clinical assessment of executive functioning and where large numbers of perseverative errors are taken to indicate an impairment of such function, requires the coordinated operation of multiple processes to maintain a sorting rule in memory, process both positive and negative feedback, inhibit the use of failed sorting rules, and infer, select and apply alternative potential sorting rules. While the WCST is often considered to be a "set-shifting" task (i.e., a task that primarily taps the executive function of set-shifting: Miyake et al., 2000), the avoidance of perseverative errors actually requires the successful coordinated functioning of several different control processes (Stuss et al., 2000). Significant research questions therefore remain concerning the operation of control processes in more complex tasks.

This paper uses a task of moderate complexity to explore the operation and potential interaction of two processes that have been held to be involved in related aspects of cognitive control. The processes are set-shifting, for which there is a well-established

literature spanning a century of work (see Monsell, 2003, and Vandierendonck et al., 2010, for reviews and, in the latter case, a theoretical synthesis), and place-keeping, a control process introduced more recently by Altmann et al. (2014) in order to account for participant performance on sequential tasks in the face of interruptions. We begin by reviewing the two processes before introducing a variant of Altmann et al.'s (2014) UNRAVEL task. We argue that performance of the task recruits both set-shifting and place-keeping processes, and present three experiments which suggest that these processes operate independently. On the basis of these results, we further argue that distinct representations of task are involved in the cognitive processes underlying set-shifting and place-keeping, and illustrate this with a computational model that simulates key behavioral effects.

1.1 Set-Shifting

"Task set" refers to the configuration of the cognitive system required to perform a specific task (see, e.g., Logan & Gordon, 2001; Schneider & Logan, 2007), including the configuration of attentional and response-selection processes, the priming or activation of learned stimulus-response associations, and so on. Thus a color-naming task set would consist of attentional orientation to color input, activating associations between visual color inputs as stimuli and the corresponding verbal names as responses, and setting a suitable response-selection threshold for verbal output, while a word-reading task set would consist of attentional orientation to visual word inputs, activation of associations between visual word inputs as stimuli and the corresponding verbalizations as responses, and perhaps a similar response-selection configuration. Set-shifting is held to be required whenever higher order processing requires changing from one task set to another (i.e., whenever task-switching is required).

While a number of empirical paradigms have been developed to explore set-shifting, it is typically evidenced by choice reaction time tasks where participants respond by categorizing blocks of stimuli in different ways on different trials. Participants take longer to respond, and are more likely to err, on trials following a change in the categorization rule compared to trials that do not require a change in the rule – the so-called switch cost. There is debate about the cognitive processes that give rise to this cost. One view is that switch costs reflect management of interference from the preceding task set on the new task set (e.g., Allport & Wiley, 2000). A competing view is that they reflect processes involved in reconfiguring the cognitive system for the new task set (e.g., Rogers & Monsell, 1995). A middle ground is advocated by Vandierendonck et al. (2010), who argue that set-shifting requires both interference management and task set reconfiguration.

One key finding in the task-switching literature is that switch costs can be reduced by forewarning participants of an impending change of set. Thus, Rogers and Monsell (1995) used a set-shifting task in which participants could anticipate the task set that would be appropriate for the forthcoming stimulus. They explored the switch cost as a function of Response-Stimulus Interval (RSI: i.e., the interval between the response on trial *n-1* and the presentation of the stimulus on trial *n*). Switch costs were lower when the RSI was high compared to when it was low, suggesting that participants were able to use the RSI to prepare for the forthcoming task. However, the reduction in the switch cost was less than the increase in RSI (e.g., increasing RSI by 150 msec from 150 msec to 300 msec led to a reduction in switch cost of approximately 40 msec; Rogers & Monsell, 1995, experiment 4), and a residual switch cost remained even with very long RSI (1200 msec). One interpretation of this result is that when a task switch is

predictable, participants may use the RSI to reconfigure the cognitive system in preparation for the upcoming task set, but that stimulus presentation is necessary to complete these reconfiguration processes (e.g., Rogers & Monsell, 1995; Meiran, 1996). Alternatively, the RSI may allow dissipation or inhibition of the activation of the previous (competing) task set, with stimulus presentation being necessary to optimize performance and minimize interference from activation of other task sets (e.g., Allport & Wiley, 2000). A third interpretation proposed by De Jong (2000) is that participants are indeed fully able to reconfigure their cognitive systems during a sufficiently long RSI, but that they do not do this on all trials. On some trials they may "fail-to-engage" in preparation. On this account the residual switch cost is an artefact of averaging RT over trials. De Jong (2000) supports this interpretation through an analysis of RT distributions.

Regardless of the precise origin of the residual switch cost, or of switch costs more generally, it is important to note for current purposes that the established view is that the processes required for task-set reconfiguration, i.e., processes related to the inhibition or disengagement of the current task set and/or the priming/engagement of the subsequent task set, must occur whenever switching from one task-set to another. That is, set-shifting is not a process that only occurs in studies that explicitly measure switch costs. It is a process, or collection of processes, that occur whenever a change of task set is required, as, for example, when it is necessary to respond to different aspects of a stimulus on successive trials.

1.2 Place-Keeping

Studies of set-shifting typically require participants to switch between two or sometimes three tasks, with the target task on any trial being cued either explicitly or implicitly by some aspect of the stimulus, such as its position on the screen. Yet many tasks, particularly real-world tasks, are temporally extended and sequential in nature – they consist of multiple steps that should be performed in order but without explicit order cues. A fundamental question therefore concerns how the cognitive system generates sequential behavior. Prior to the cognitive revolution is had been argued that such behavior was the result of stimulus-response S-R chains (e.g., Washburn, 1916; Watson, 1920; cited in Lashley, 1951), with one response in a sequence serving as the stimulus for the next S-R link. Lashley (1951) argued against this account on the grounds that the elements of sequential behaviors (both in the domains of language and skilled movement) have super-ordinate structure: R_1 is not invariably followed by R_2. Rather, R_1 might be followed by R_2 in one situation but by R_3 in another. Lashley's alternative proposal was that sequential behavior was governed by higher-order task representations which excite or partially activate elements of the sequence, with sequential order of those elements achieved through dynamic processes (the details of which he did not specify) operating on the partially activated elements.

More recently, Altmann and colleagues (Altmann et al., 2014; Altman & Trafton, 2015) have proposed a process or mechanism of place-keeping, which, they argue is involved in maintaining one's position in temporally extended sequential tasks. To date, Altmann and colleagues have investigated place-keeping by exploring the effects of interruptions on participants' performance of a sequential task (i.e., on a task consisting of multiple steps that are required to be performed in a strict order). Errors in behavior following an interruption have frequently been used to inform theories of task maintenance (e.g., Byrne & Bovair, 1997; Botvinick & Bylsma, 2005; Altmann & Trafton, 2007; Li et al., 2008), but even in the absence of interruptions, performance of sequential tasks is prone

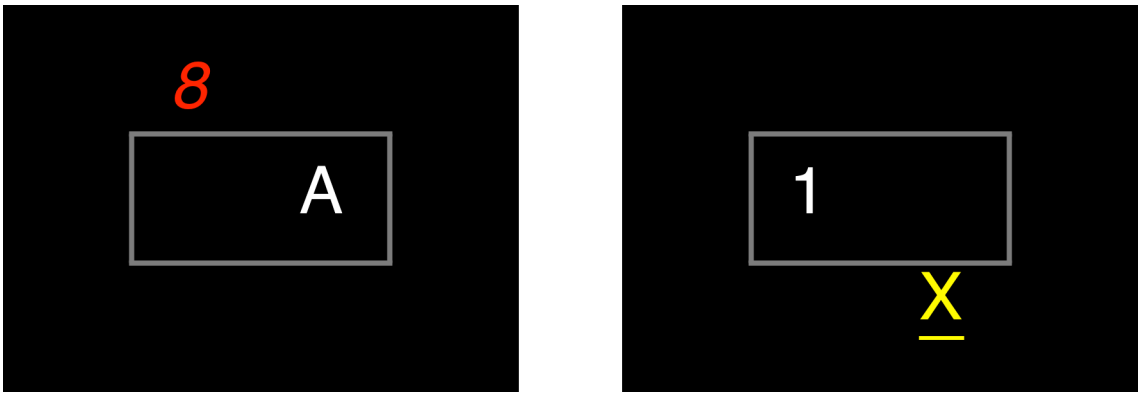
to both perseverative errors (where steps are inappropriately repeated) and anticipatory errors (where steps are left out).[2] Altmann et al. (2014) found that interruptions, even of just a few seconds, increase the probability of perseverative errors more than the probability of anticipatory errors. The authors interpret their findings within an information-processing model in which sequential processing relies upon associative links from the representation of the current step to the representation of the next step (i.e., a form of context-specific chaining), combined with activation decay to discourage reselection of recently completed steps (see Altman & Trafton, 2015).

Hambrick and Altmann (2015) elaborate on these findings by showing in an individual differences study that place-keeping ability (measured in terms of sequential accuracy on the interrupted UNRAVEL task, as described below) correlates with fluid intelligence (as measured with Raven's advanced progressive matrices), but not with working memory capacity, set-shifting or multi-tasking ability. They further argue that place-keeping is itself an executive function, i.e., one of a "suite of mental operations that coordinate and supervise other mental operations in the service of performing some task" (p. 104), and relate the function to more general mechanisms responsible for goal-management. On this view it is of similar theoretical importance to more widely accepted executive functions, such as set-shifting as described above. Critically, place-keeping is argued by Altmann and colleagues to be pervasive. That is, while their empirical investigations of place-keeping have been focused on the effects of interruption within a specific sequential task, the mechanisms that support place-keeping are held to operate in any complex task where steps must be completed in a prescribed order.

1.3 The UNRAVEL Task

The UNRAVEL task (Altmann et al., 2014; Altmann & Trafton, 2015; Hambrick & Atlmann, 2015; Altmann & Hambrick, 2017) is a sequential binary choice response task in which successive steps involve responding to different dimensions of a compound stimulus. The order of the steps (i.e., the dimension for each successive response) is encoded in the acronym "UNRAVEL", which serves as mnemonic. Two potential stimuli from the task are shown in figure 1. On the first step, participants are instructed to respond (by typing with their preferred fingers on a standard computer keyboard) with the letter 'U' if the stimulus contains an underlined character or 'I' if it contains an italic character. (Exactly one character in each stimulus is either underlined or italic.) On the second step, participants are instructed to respond with the letter 'N' if the letter in the stimulus is near the beginning of the alphabet and 'F' if it is far from the beginning. (Letters in the stimulus are either 'A', 'B', 'U' or 'X', so again there is no ambiguity in the response.) The procedure continues, with the possible responses on each step given by the rules in figure 1 and the step determined by cycling through the seven letters of "UNRAVEL" (i.e., after the 'L' step, participants are instructed to return to the 'U' step). Critically, the response options on each step are different, so it is possible to infer from a participant's response their place in the UNRAVEL sequence.

[2] Whether anticipatory or perseverative errors dominate appears to depend on the specific sequential task (cf. Altmann et al., 2014, versus Trafton et al., 2011).



Response Options		Response Rules
U	I	Is a character Underlined or in Italic font?
N	F	Is the letter Near to or Far from the beginning of the alphabet?
R	Y	Is a character in Red or Yellow?
A	B	Is a character Above the box or Below the box?
V	C	Is the letter a Vowel or a Consonant?
E	O	Is the digit Even or Odd?
L	M	Is the digit Less than five or More than five?

Figure 1: Two sample stimuli from the UNRAVEL task (upper panel) and the seven response rules (lower panel). Adapted from Altmann et al. (2014).

1.4 Set-Shifting and Place-Keeping in the UNRAVEL Task

The UNRAVEL task was initially developed as a means of exploring error following interruption on a sequential task. Thus, Altmann, Trafton and colleagues (Altmann et al., 2014; Altmann & Trafton, 2015) compared the types of errors made by participants when interrupted at random points in the UNRAVEL sequence. On trials that were not preceded by an interruption, participants occasionally left steps out (i.e., they made anticipation errors). More rarely they repeated steps (i.e., they made perseverative errors). On trials following a brief interruption (of 2 to 3 seconds), however, the rate of anticipatory errors increased slightly (from approximately 1.0% of trials to 1.5% of trials), while the rate of perseverative errors increased more markedly (from less than 0.5% to almost 3.0% of trials), with longer interruptions resulting in even more perseverative errors.

As noted above, Altmann et al. (2014) interpret their results in terms of a cognitive operation of "place-keeping" that is held to be involved in maintaining one's position within a sequential task. However, the basic (uninterrupted) UNRAVEL task also shares many key features with tasks commonly used to investigate set-shifting. As with the tasks reviewed by Monsell (2003), for example, stimuli are multidimensional and on

different trials participants are required to respond to different dimensions, switching the dimension of the response (and hence the task set) on successive trials. In other words, the basic UNRAVEL task is a kind of set-shifting task in which the participant must shift set on each trial, using an internally maintained record of his/her position in the UNRAVEL sequence (which must be updated between each step) to determine the task set appropriate for each successive stimulus.

The UNRAVEL task does not include repeat trials with the same task set, and so it does not afford calculation of, for example, switch costs. However, if the basic UNRAVEL task does recruit set-shifting processes, then one would anticipate that variables known to affect set-shifting performance should also affect performance on the UNRAVEL task. As discussed above, one such variable that has been extensively studied is the response-stimulus interval. Given that the required task set on successive UNRAVEL trials is deterministic and known to the participant, increasing the response-stimulus interval between UNRAVEL trials should, on the assumption that the UNRAVEL task requires set-shifting, result in faster response times.

A second variable of interest in the (interrupted) UNRAVEL task is the temporal duration of interruptions. In the experiments of Altmann et al. (2014), participants were occasionally interrupted by a secondary task between trials of the UNRAVEL task. The secondary task involved using the keyboard to copy either 4 (experiment 1) or 2 (experiment 2) characters that appeared on screen. Unsurprisingly, temporally longer durations led to more sequence errors. Interruption length is therefore a key variable that appears to affect the effectiveness of place-keeping processes on the interrupted UNRAVEL task.

1.5 Overview of the Argument and Structure of the Paper

Given the above context, the experiments reported below vary response-stimulus interval and interruption length within the UNRAVEL task in order to determine whether the control processes that support set-shifting and place-keeping are functionally independent. If the processes are independent, then introducing (or increasing) a response-stimulus interval within the UNRAVEL task should result in decreased response times on UNRAVEL trials, but should not affect measures of place-keeping (i.e., sequence error rates). We consider the former (i.e., decreased RT) to be a "fingerprint" of the processes supporting set-shifting. Applying analogous logic, variation of interruption length should affect sequence error rates (in accordance with previous studies of the UNRAVEL task, and indicating the operation of place-keeping mechanisms), but should not affect measures relating to set-shifting. Alternatively, if participants use intervals between UNRAVEL trials to prepare for the next step, whether that preparation involves activation of the forthcoming set or inhibition/dissipation of the prior set (as the set-shifting literature would suggest), and the results of that preparation are accessible to or shared by processes that support place-keeping, then an interruption after the response-stimulus interval but prior to actual presentation of the stimulus should be less disruptive with longer response-stimulus intervals (where greater task-set reconfiguration or dissipation may take place) than with shorter or no response-stimulus intervals (where more limited or no task-set reconfiguration/dissipation could take place). Regardless of this issue, if place-keeping and set-shifting are task-independent control functions, then the same pattern of results should hold regardless of the interrupting task. We therefore report three experiments, using different interrupting tasks, that explore the effects of response-stimulus interval and interruption length on a range of dependent measures. The results indicate that the

mechanisms underlying set-shifting and place-keeping are separable. In the discussion we consider implications of these findings for theoretical and computational accounts of the two control functions, and extend an existing computational model of set-shifting with place-keeping mechanisms in order to provide a theoretically novel account of our results.

2 Experiment 1

Experiment 1 sought to establish the independence or otherwise of place-keeping and set-shifting by replicating the Altmann et al. (2014) study but manipulating two variables: the response-stimulus interval within the UNRAVEL task and the length of the interruption. If successive steps on the UNRAVEL task involve set-shifting as argued in the introduction, and if place-keeping and set-shifting are independent, then manipulation of the response-stimulus interval within the UNRAVEL task should affect response times on that task (with longer response-stimulus intervals allowing greater preparation of the upcoming subtask, and hence resulting in shorter response times), but such a manipulation should have no effect on performance of the interrupting task or on resumption of the UNRAVEL task. Conversely, interruption length should affect the likelihood of a sequence error occurring on resumption of the UNRAVEL task (with longer interruptions resulting in more sequence errors) and this should not be modulated by response-stimulus interval. The alternative possibility is that the greater task-set preparation afforded by longer response-stimulus intervals will improve place-keeping and hence result in lower post-interruption error rates on the UNRAVEL task when the response-stimulus interval is longer.

2.1 Method

2.1.1 Participants

Forty-five adult participants from the Birkbeck student community completed the study in exchange for course credit. Data from five participants were excluded because of equipment failure. The final sample therefore comprised 40 participants (25 female, mean age 27.3 years; 15 male, mean age 24.2 years).

2.1.2 Design

The experiment involved manipulation of two within-subjects factors: response-stimulus interval (at three levels: 0 msec, 250 msec, and 500 msec) and interruption length (short, medium, long). The dependent measures were accuracy and response time on each task (i.e., on the UNRAVEL task and on the interrupting task, as described below) and number and type of errors on baseline trials of the UNRAVEL task (i.e., those not preceded by an interruption) and post-interruption trials of the UNRAVEL task.

2.1.3 Interrupting Task: The Text-Copying Task

In experiment 1 the interrupting task was a text-copying task similar to that used by Altmann et al. (2014). Thus, between some trials of the UNRAVEL task a screen appeared asking participants to use the keyboard to copy a sequence of letters. In contrast to Altmann et al., the number of to-be-copied letters varied on interrupting trials (with three levels: 4, 8 or 12, for short, medium and long interruptions respectively), and the letters were drawn from those that were not in the UNRAVEL response set (i.e., there were drawn from DGHJKPQSTWXZ). Participants were free to use their preferred fingers for typing their response in the text-copying task. On

completion of the task participants were instructed to press the return key, which was only enabled when the correct number of letters had been typed. This led to immediate continuation of the UNRAVEL task from the point at which the interruption occurred. Note that participants were not required to achieve perfect accuracy on the text-copying task, though accurate copying was stressed, and all participants achieved high levels of accuracy on the task. (See the results subsection.)

2.1.4 Procedure

Participants were first given an in-depth verbal introduction to the UNRAVEL task. In this introduction the acronym was highlighted, emphasizing each letter step-by-step together with its corresponding pair of response options. Once the task had been explained, participants were directed to a written summary of the instructions for each step, which was posted on the wall in their line of sight. This summary remained visible throughout the task. No instructions were given about typing responses with specific fingers (i.e., participants were free to use their individual typing preferences). Participants then completed two practice blocks of the task (administered on a standard PC with 21" monitor and 16:9 aspect ratio) to ensure that they understood the task.

The first practice block consisted of 56 UNRAVEL trials (i.e., 8 repetitions of the full UNRAVEL sequence) and no interruption trials. Interruption trials, where participants were presented with a screen asking them to copy a string of letters as described above, were introduced on the second practice block. In this block there were two 4 letter, two 8 letter and two 12 letter interruptions, distributed randomly throughout the UNRAVEL sequence according to a translated exponential distribution with minimum separation of 3 UNRAVEL trials and a mean separation of 6 UNRAVEL trials between interruptions, as per the specification of Altmann et al. (2014). The mean number of UNRAVEL trials in this second practice block was 42, but the number of trials per participant varied depending on the number of trials between interruptions, which itself varied as described above. An experimenter remained present during the practice blocks to give assistance if needed. Moreover, during the practice blocks the program that administered the task provided feedback to participants on a trial-by-trial basis if they made an error on either task, indicating the correct response. In the case of the UNRAVEL task, and in this second practice block only, the program also indicated from which step the participant should resume. The response-stimulus interval (i.e., the interval between UNRAVEL trial responses and the presentation of the subsequent stimulus, whether it be another UNRAVEL trial or the interruption task) in the practice blocks was fixed within participants but varied between participants, being either 0 msec, 250 msec or 500 msec. Note that, in all cases, completion of the interrupting task led immediately to resumption of the UNRAVEL task. That is, in all cases there was no interval between the participant entering their response to the interrupting task and presentation of the next UNRAVEL stimulus.

Participants then completed 6 experimental blocks (without the experimenter present). Each experimental block was structured as in the second practice block but with three short, three medium and three long interruptions (and still a mean of six UNRAVEL trials between interruptions), resulting in a mean of approximately 60 UNRAVEL trials per block. The response-stimulus interval was the same on the first and fourth experimental blocks as in the practice blocks. On the second and fifth experimental blocks (and likewise the third and sixth experimental blocks) it took on one of the other possible values, with the permutation of response-stimulus intervals for experimental blocks 1, 2 and 3, randomized across participants. Feedback on individual trials was not

given within the experimental blocks. Instead, accuracy feedback was provided on screen at the end of each block (e.g., "That completes block 1. You scored 95% on this block."). Each participant was individually tested in one session lasting between 45 and 60 minutes.

As in previous work with the UNRAVEL task, a different compound stimulus was presented on each trial, with successive stimuli differing in the digit, the letter, the position of the element outside of the box, the position or color of the colored element, and the position or font decoration of the italic/underlined element. Moreover, participants potentially had unlimited time to respond on each trial (i.e., trials did not time-out).

Prior to embarking on data collection the experimental procedure was reviewed and approved by the departmental ethics committee.

2.2 Results

2.2.1 UNRAVEL Accuracy and Response Time

One would anticipate that speed and accuracy on the UNRAVEL task will vary as a function of trial type. For example, one would anticipate faster and more accurate responses on 'R' trials (discriminating red or yellow) than 'V' trials (discriminating whether the letter is a vowel or consonant), simply because the former is a more practiced discrimination. Such an analysis would not be informative with respect to our hypotheses. We therefore pool the data from all trial types and focus first on UNRAVEL accuracy and response times on baseline trials (i.e., on trials not immediately preceded by an interruption) as a function of response-stimulus interval.

UNRAVEL trials with response time of greater than 20 seconds (less than 0.1% of trials) were excluded from analysis, as were those immediately following an interruption. On the remaining trials, accuracy was generally high (see figure 2, left), resulting in a negatively skewed distribution. A one-way within-subjects ANOVA on the accuracy data (arcsine transformed, to reduce skew) revealed no effect of response-stimulus interval ($F(2, 78) = 0.009$, p = 0.991, $\eta_p^2 = 0.000$). In contrast, response time on correct trials decreased with response-stimulus interval, from a mean of 2610 msec to a mean of 2295 msec as response-stimulus interval increased from 0 msec to 500 msec (see figure 2, right), and a one-way within-subjects ANOVA revealed that the effect of response-stimulus interval on response time was significant ($F(2, 78) = 15.685$, $p < 0.001$, $\eta_p^2 = 0.287$). Follow-up t-tests indicated that all pairwise differences were significant (0 msec versus 250 msec: $t(39) = 2.847$, $p = 0.007$; 250 msec versus 500 msec: $t(39) = 3.258$, $p = 0.002$; two-tailed probabilities in both cases).

Our interpretation of the UNRAVEL task as a switching task with interruptions implies that response time on post-interruption trials (where participants must recover their position in the UNRAVEL sequence) should be longer than on baseline trials (where position in the UNRAVEL sequence is more immediately available based on the previous trial). At the same time, since there was no response-stimulus interval between termination of an interruption and presentation of the first post-interruption UNRAVEL trial, baseline and post-interruption response times can only be directly compared in the 0 msec RSI blocks. A one-within ANOVA on the RT for correct trials from these blocks, comparing across four levels of interruption length (none, short, medium and long), revealed a significant effect of interruption length ($F(3, 117) = 4.126$, $p = 0.008$,

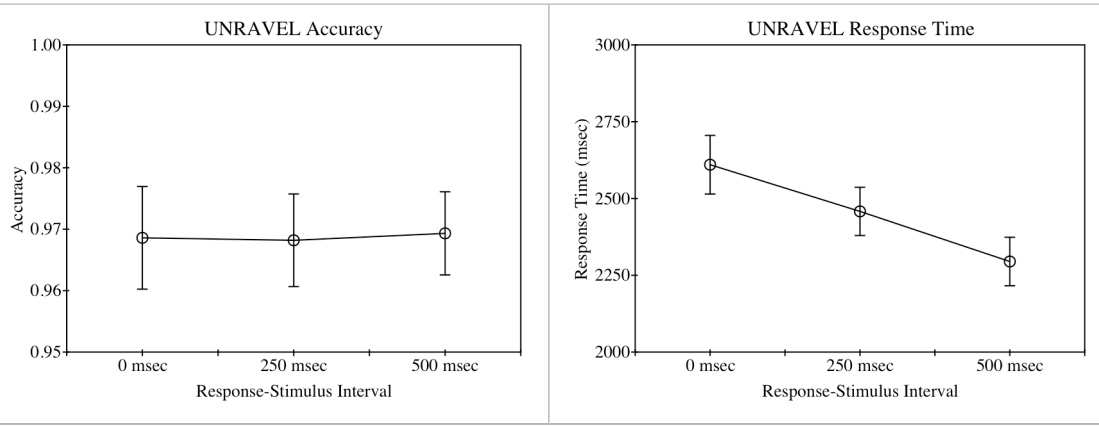


Figure 2: UNRAVEL accuracy (left) and response time (right), for trials not preceded by an interruption, as a function of response-stimulus interval in experiment 1. Error bars indicate one standard error.

η_p^2 = 0.096). While the rank order of mean response times reflected interruption length (2610 msec, 2629 msec, 2861 msec, 3094 msec for none, short, medium and long interruptions), the only statistically significant pairwise difference was that between RT following long interruptions compared with baseline trials (i.e., 3094 msec versus 2610 msec: $t(39) = 2.885$, $p = 0.006$, two-tailed).

With regard to potential effects of response-stimulus interval on post-interruption response time our interpretation does not make clear predictions given that the RSI preceded but did not follow interruptions. Arguably, such effects might be attenuated (compared to baseline trials), given that any preparation prior to an interruption will wholly or partially dissipate during the completion of the interrupting task. A two-way within-subjects ANOVA on response time on correct post-interruption UNRAVEL trials was performed to test whether the effect of response-stimulus interval on baseline UNRAVEL trials persisted in post-interruption trials. While mean response times were slower for short response-stimulus intervals than long response-stimulus intervals (2861 msec at 0 msec RSI versus 2686 msec at 500 msec RSI), the effect was not significant (main effect of response-stimulus interval: $F(2, 78) = 1.735$, $p = 0.183$, $\eta_p^2 = 0.043$). However, there was (as would be anticipated from the preceding analysis) an effect of interruption length ($F(1.551, 60.488) = 10.038$, $p = 0.001$, $\eta_p^2 = 0.205$) on post-interruption response time, with the longest interruptions leading to significantly longer response times than the shortest interruptions (2967 msec versus 2584 msec; $t(39) = 3.775$, $p = 0.001$). Response-stimulus interval and interruption length did not interact in their effects on post-interruption response time ($F(4, 156) = 0.934$, $p = 0.446$, $\eta_p^2 = 0.023$).

2.2.2 The Text-Copying Task

In the analysis of performance on the interrupting task responses of more than 4 seconds per character were removed. (Only one response was excluded based on this criterion). Accuracy (i.e., the proportion of characters copied correctly) and response times, as a function of response-stimulus interval and interruption length, of the remaining data are shown in figure 3. As the figure shows, accuracy was high in all conditions, with no apparent effect of condition. The accuracy data was arcsine transformed to reduce skew due to the ceiling effect. A two-way ANOVA on the resultant data revealed no

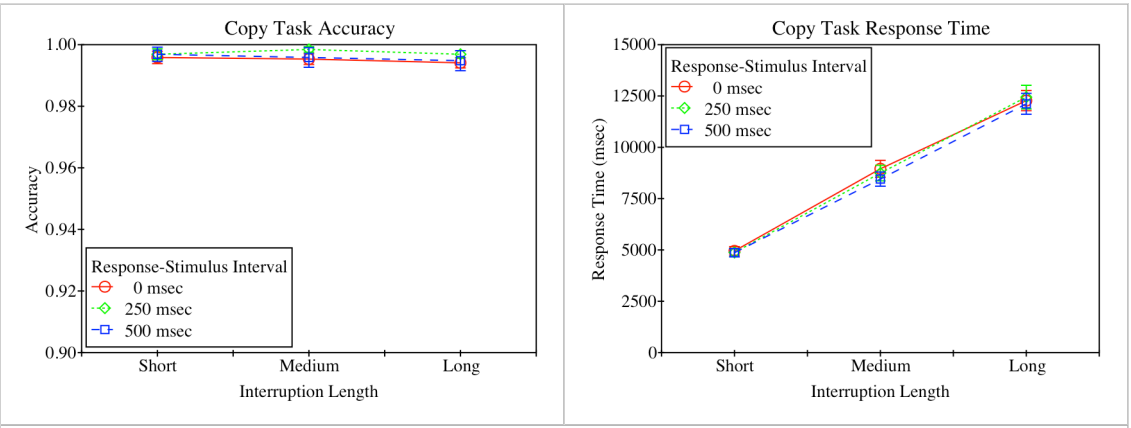


Figure 3: Interruption task accuracy (left) and response time (right) as a function of response-stimulus interval and interruption length in experiment 1. Error bars indicate one standard error.

significant effect of response-stimulus interval ($F(2, 78) = 1.330$, $p = 0.270$, $\eta_p^2 = 0.033$), no significant effect of interruption length ($F(2, 78) = 1.919$, $p = 0.154$, $\eta_p^2 = 0.047$), and no significant interaction ($F(4, 156) = 0.225$, $p = 0.924$, $\eta_p^2 = 0.006$).

A similar ANOVA was performed on response times for the interrupting task. The analysis revealed a significant effect of interruption length ($F(1.169, 45.584) = 392.812$, $p < 0.001$, $\eta_p^2 = 0.910$), as would be expected, but no significant effect of response-stimulus interval ($F(2, 78) = 2.430$, $p = 0.095$, $\eta_p^2 = 0.059$) and no interaction between response-stimulus interval and interruption length ($F(2.840, 110.769) = 1.069$, $p = 0.363$, $\eta_p^2 = 0.027$).[3] Planned comparisons revealed, as expected that 12 letter interruptions took significantly longer than 8 letter interruptions at each level of response-stimulus interval ($t(39) \geq 11.455$, $p < 0.001$, in all cases) which in turn took significantly longer than 4 letter interruptions ($t(39) \geq 14.259$, $p < 0.001$, in all cases). This analysis confirms that the interruption length manipulation was effective in altering the time participants spent performing the interrupting task and validates the following analyses of post-interruption responses as a function of interruption length.

2.2.3 Sequence Errors on the UNRAVEL Task

Recall that a key feature of the UNRAVEL task is that the response options on each step are different. Consequently one can infer, from a participant's response, the step to which the participant was responding. Critically, one can therefore detect and categorize sequence errors. Following Altmann et al. (2014), we encode such errors in terms of whether they are anticipatory (i.e., skipping forward) or perseverative (i.e., repeating a recent step), and in terms of the difference between the expected and actual step. More specifically, if a participant skips a step (e.g., responding with a 'R' or 'Y' when on the N step), we encode the response as an "A1" (anticipate one step) sequence error. If, on the other hand, the participant repeats the step before last (e.g., responding with an 'N' or 'F' when on the A step), we encode the response as a "P2" (perseverate two steps) sequence error.

[3] Sphericity was violated for all analyses involving interruption length (both here and in experiments 2 and 3), and so Greenhouse-Geisser corrections to degrees of freedom were applied in all cases.

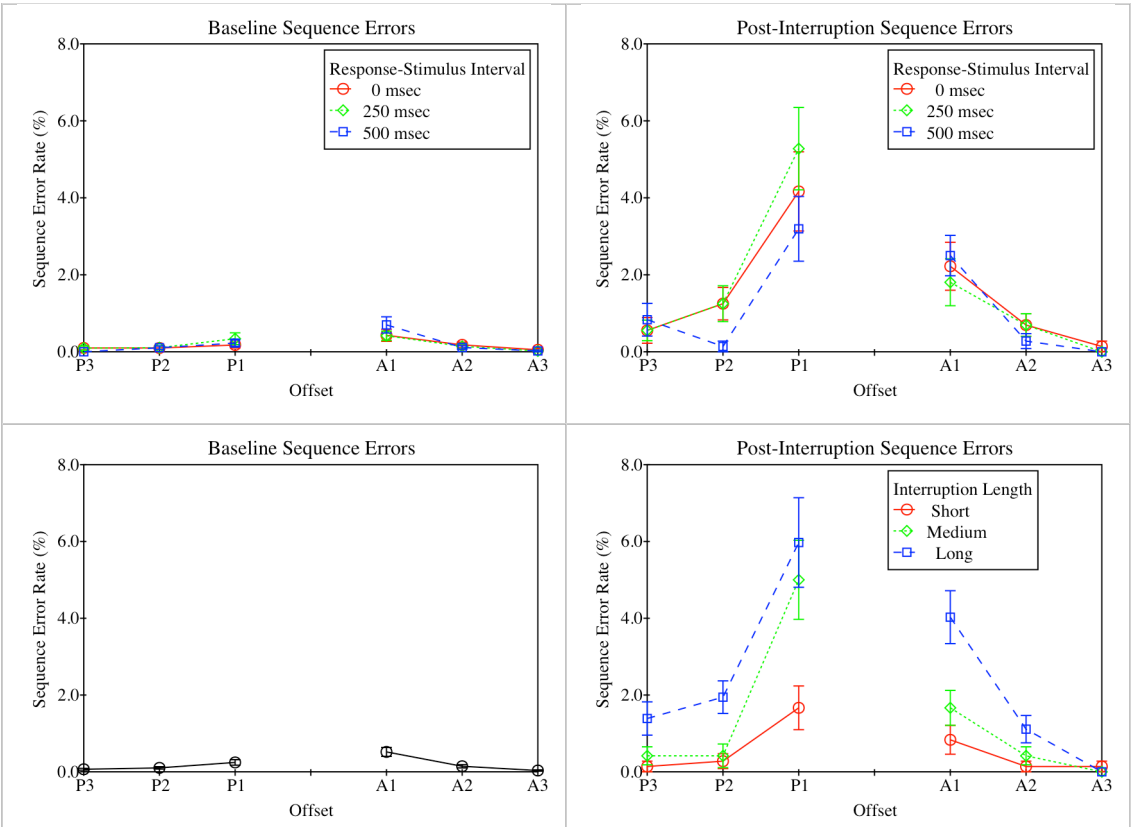


Figure 4: Sequence error rates on baseline UNRAVEL trials (i.e., those not preceded by an interruption) (left) and on post-interruption UNRAVEL trials (right) in experiment 1. The upper panels show the data by response-stimulus interval (pooled over interruption duration in the upper right panel), while the lower panels show the data pooled over response-stimulus interval. In all cases error bars indicate one standard error.

Figure 4 shows the proportion of errors of each type for baseline (i.e., non-interrupted) UNRAVEL trials (left) and post-interruption UNRAVEL trials (right, showing each level of interruption length) in experiment 1. Given the experimental design, each participant completed, for each level of response-stimulus interval, approximately 102 baseline trials and exactly 18 trials following an interruption (6 trials following a short interruption, 6 trials following a medium interruption, and 6 trials following a long interruption). Thus, each participant had 102 opportunities to err on baseline trials in each of the three (within-subjects) cells of the experimental design but only 6 to err following an interruption in each of the nine (within-subjects) cells. The proportion of post-interruption errors, particularly when considered within a single cell of the experimental design, is therefore a coarse measure. Moreover many participants performed without error in at least one experimental cell. Therefore, for clarity, the upper panels of figure 4 show the data for each response-stimulus interval condition (but pooled over interruption duration where applicable) while the lower panels show the data pooled over response-stimulus interval.

There are numerous analyses that might be performed on the sequence error data. One possibly surprising finding reported by Altmann et al. (2014) concerns the relative rates of perseverative and anticipatory errors in baseline (i.e., UNRAVEL trials not preceded by an interruption) as opposed to post-interruption trials. As with Altmann et al.'s study,

anticipatory errors were observed to be more common than perseverative errors following baseline trials, but the opposite was observed in post-interruption trials. Elaborating on Altmann et al.'s results, a comparison of P1 versus A1 error rates (log transformed, to reduce skew) as a function of interruption level (none, short, medium, long), using a 2 (error type) by 4 (interruption level) ANOVA, found no main effect of error type ($F(1, 39) = 2.953$, $p = 0.094$, $\eta_p^2 = 0.070$), but a main effect of interruption level ($F(2.355, 91.833) = 19.676$, $p < 0.001$, $\eta_p^2 = 0.335$), and a significant interaction ($F(2.449, 95.526) = 3.258$, $p = 0.033$, $\eta_p^2 = 0.077$), suggesting that the rates of perseverative and anticipatory errors varied as a function of interruption level. Follow-up related t-tests on P1 versus A1 error rates revealed a significant difference in favor of A1 errors on baseline trials ($t(39) = -2.603$, $p = 0.013$, two-tailed), no significant difference following short interruptions ($t(39) = 1.308$, $p = 0.199$, two-tailed), and a significant difference in favor of P1 errors following medium interruptions ($t(39) = 2.631$, $p = 0.012$, two-tailed). While P1 errors were relatively more frequent than A1 errors following long interruptions, the comparison did not approach significance ($t(39) = 0.412$, $p = 0.683$, two-tailed).

The preceding analysis collapses across response-stimulus interval on the UNRAVEL task, but a second key question concerns whether response-stimulus interval on the UNRAVEL task affects the rate of each type of error, either on baseline trials or on trials immediately following an interruption. A differential effect of response-stimulus interval across interruption length would suggest some form of shared mechanism underlying the cognitive processes that support set-shifting and place-keeping. In order to explore this question a 3 (response-stimulus interval) × 4 (interruption level: none, short, medium, long) multivariate analyses of variance (MANOVA) was performed on the (log transformed) proportions of each type of error (i.e., with six dependent measures: A1 errors, A2 errors, A3 errors, P1 errors, P2 errors and P3 errors). This analysis revealed no significant interaction between response-stimulus interval and interruption length (Pillai's Trace = 0.912, $F(33, 7) = 2.207$, $p = 0.139$, $\eta_p^2 = 0.912$) and no significant effect of response-stimulus interval (Pillai's Trace = 0.465, $F(12, 28) = 2.024$, $p = 0.061$, $\eta_p^2 = 0.465$). In contrast to the absence of effects related to response-stimulus interval, the effect of interruption length was significant (Pillai's Trace = 0.706, $F(17, 23) = 3.256$, $p = 0.005$, $\eta_p^2 = 0.706$).

In considering the results of this MANOVA it should be stressed that the data are heavily skewed as errors are relatively rare, particularly following base-line trials.[4] Consequently the results must be interpreted with caution. While it appears that interruption length does, and response-stimulus interval does not, affect the rate of each type of error, additional evidence is necessary to make a more convincing case. In particular, to test for a null effect of response-stimulus interval (or its interaction with interruption length) on sequence errors, a series of Bayesian repeated measures ANOVAs were conducted on the sequence error data (untransformed proportions) using the JASP package (JASP Team, 2018; Version 0.8.5.1). Bayes factors for inclusion of each of the three potential model factors (RSI, interruption length (IL), and their interaction) for each error type (P3, P2, P1, A1, A2, A3) are shown in table 1. With the exception of A2 and A3 errors (which were rare), the data provide strong evidence ($BF_{Inc} >> 1$) for models with a single factor (interruption length). In contrast, the results

[4] Reanalysis of the data excluding the baseline trials, i.e., via a 3 × 3 MANOVA, reduces the violation of homogeneity of variance but yields qualitatively equivalent results.

	P3	P2	P1	A1	A2	A3
RSI	0.021	0.629	0.069	0.029	0.039	0.047
IL	34.563	356.533	3.6×10^5	4.9×10^5	1.158	0.019
RSI × IL	0.002	0.880	0.015	0.002	0.024	2.4×10^{-4}

Table 1: Bayes Factors (BF_{Inc}) for inclusion of RSI, IL (interruption length) and their interaction in models of each sequence error type (experiment 1). Note that with the exception of A2 and A3 errors (which were rare), the data provide strong evidence for models ($BF_{Inc} >> 1$) with a single factor (interruption length). In contrast, the results generally provide evidence against models ($BF_{Inc} << 1$) that include RSI or the RSI × IL interaction term.

generally provide evidence against ($BF_{Inc} << 1$) models that include RSI or the RSI × IL interaction term.

2.3 Discussion

The results of experiment 1 a) support the interpretation of the UNRAVEL task as involving the same cognitive processes as involved in more typical set-shifting tasks, b) extend the results of Altmann et al. (2014) concerning place-keeping and the relative proportions of perseverative and anticipatory errors in baseline performance and following an interruption, and c) suggest that cognitive processes that support set-shifting and place-keeping operate independently.

First, response time on the UNRAVEL task was faster with larger response-stimulus interval. This is consistent with an interpretation of UNRAVEL performance whereby participants are able to use the response-stimulus interval to prepare for the forthcoming task (either by activating components of the forthcoming task-set or by inhibiting those of the previous task-set), as in standard task-switching procedures where the response-stimulus interval is varied (e.g., Rogers & Monsell, 1995). Note however that increasing the response-stimulus interval from 0 msec to 250 msec lead to a reduction in response time of 142 msec, while increasing it from 250 msec to 500 msec led to a further reduction of 167 msec. In other words, when preparation time was available, the reduction in response time was somewhat less than the increase in preparation time, but further extending the preparation time resulted in a significantly greater reduction in response time. Moreover the reduction in response time did not appear to have been the result of a speed-accuracy tradeoff, as accuracy was not poorer in the 500 msec RSI condition (when responses were relatively fast) than in the 0 msec RSI condition (when responses were relatively slow).[5] These results are all as would be expected based on prior work within more standard task-switching paradigms and support the claim that set-shifting processes are involved in configuring the cognitive system (through active

[5] Recall that participants completed two blocks of UNRAVEL trials at each level of response-stimulus interval. Further analysis, reported in the supplementary materials, found that UNRAVEL response time was faster on the second block for each level of response-stimulus interval, as would be expected from learning, but that the speed-up was independent of response-stimulus interval. That is, the reduction in response time was affected by response-stimulus interval but not practice, further supporting our conjecture that task-switching is an integral process within the UNRAVEL task. This independence of response-stimulus interval and practice on UNRAVEL response time held for all three studies reported here.

excitation of the upcoming task-set or through inhibition/dissipation of the previous task-set) between steps of the UNRAVEL task.

Second, interruption length was, unsurprisingly, proportional to the number of letters copied in the text-copying task. That is, "larger" interruptions did indeed take longer. While this is to be expected, it licenses analysis of sequence error behavior as a function of interruption length. This analysis demonstrates that: a) in trials not preceded by an interruption, anticipatory sequence errors are more common than perseverative sequence errors, but b) perseverative sequence errors increase in relative frequency following an interruption, and this increase is monotonically related to the duration of the interruption. These results extend those of Altmann et al. (2014) (who considered 2 letter and 4 letter interruptions in separate experiments), by showing in a single study that the rates of both anticipatory and perseverative errors increase with the duration of the interruption, though the rate of perseverative errors appears to increase more than that of anticipatory errors with longer interruptions. The increase in the rate of each error type can therefore not be attributed to between-study differences (e.g., subtle differences in task instruction or participant characteristics, as might be argued for the results of Altmann et al., 2014), but reflect within-subject differences in task-related processing during interruptions of different lengths.

Third, while response-stimulus interval affected performance on the UNRAVEL task, it had no discernable effect on performance of the interrupting task (text-copying), either in terms of response time or accuracy. More critically, response-stimulus interval did not affect sequence error rates, either on baseline trials (i.e., those trials not preceded by an interruption) or on post-interruption trials. Thus, the effects of response-stimulus interval and size of interruption appear to be orthogonal. This suggests that the information-processing mechanisms responsible for set-shifting and those responsible for place-keeping are independent. More specifically, it appears that task-position maintenance during an interruption operates on a different representation of task-set than the mechanism that achieves task-set activation within the sequential task. As argued in the introduction, this is not a necessary result.

While increasing the RSI from 0 msec to 500 msec appears to allow preparation of the next step, it does not alter the relative or absolute number of anticipatory or perseverative errors on such steps. This suggests a stage-like approach to task-set processing within the UNRAVEL task, where the current step is first retrieved and then prepared. Altering the response-stimulus interval affects preparation but not retrieval of the current step. Interruptions, by contrast, disrupt retrieval, specifically favoring the production of perseverative errors, with all levels of interruption having a more dramatic effect on perseverative errors than on anticipatory errors.

A final issue that is worthy of note arises from a direct comparison of our results with those of Altmann et al. (2014). While the pattern of sequence error effects shown in figure 4 above replicates that of the earlier studies (cf. figure 4 of Altmann et al., 2014), there are quantitative differences. Thus, the baseline sequence error rate was higher in Altmann et al.'s studies (e.g., A1 errors accounted for 1.0% of baseline responses in their experiment 1, and 1.1% in their experiment 2, whereas in our study the figure was 0.5%), as was the rate of sequence errors following a four-letter interruption (1.7% compared with 0.8%). These quantitative differences may be the result of participant characteristics or subtle procedural differences between the studies. It is reassuring,

however, that whatever the origin of these differences might be, they do not alter the qualitative effects of the independent variables.

3 Experiment 2

Experiment 2 was designed to replicate the findings of experiment 1 using a different interrupting task, thereby supporting the claim that the effects found in experiment 1 reflect effects attributable to control processes rather than to the specific interrupting task. That is, if place-keeping and set-shifting are effected through independent processes or processes operating on separate representations, then the effects of interruption length and response-stimulus interval should hold regardless of the specific interrupting task. Moreover, given that the argument for the independence of the underlying processes is driven by the absence of an interaction between the variables (i.e., a null effect, albeit one in the presence of other effects), then failure to find any interaction on the key dependent measures in a second study will strengthen the conclusions of experiment 1.

3.1 Method

3.1.1 Participants

Forty-three adult participants were recruited from Birkbeck's volunteer panel and paid £7.50 to complete the study. Data from three participants were excluded because of equipment failure. The final sample therefore comprised 40 participants (26 female, mean age 27.8 years; 14 male, mean age 36.1 years).

3.1.2 Design

The design was the same as for experiment 1.

3.1.3 Interrupting Task: Dot Counting

In experiment 2 the interrupting task was a dot-counting task. For this task, the screen showed a rectangular box (450 pixels wide, 350 pixels high) containing up to 13 red dots, with each dot being 10 pixels in diameter and positioned at random and without overlap within the box. The box was centered on the screen, which was a standard 1680 by 1050 pixel 21" monitor. Participants were asked to count the number of dots and use the keyboard to enter their answer. No instructions or constraints were given on use of specific fingers for typing the response. The primary UNRAVEL task resumed as soon as the participant pressed the return key (i.e., as in experiment 1, there was no interval between responding to the interruption and presentation of the subsequent UNRAVEL stimulus).

3.1.4 Procedure

The procedure was identical to experiment 1 except that the interrupting task (the text copying task of experiment 1) was replaced with the dot counting task described above. For short interruptions there were 3, 4 or 5 dots. For medium interruptions there were 7, 8 or 9 dots. For long interruptions there were 11, 12 or 13 dots. In all cases participants were required to give a numeric answer, but that answer did not need to be correct for the resumption of the UNRAVEL task. As in experiment 1, participants were given practice on the interrupting task prior to the main experiment (again as part of the second block of practice trials) and feedback was given on the practice trials but not on the experimental trials. Also as with experiment 1, the experiment was approved by the departmental ethics committee prior to embarking on data collection.

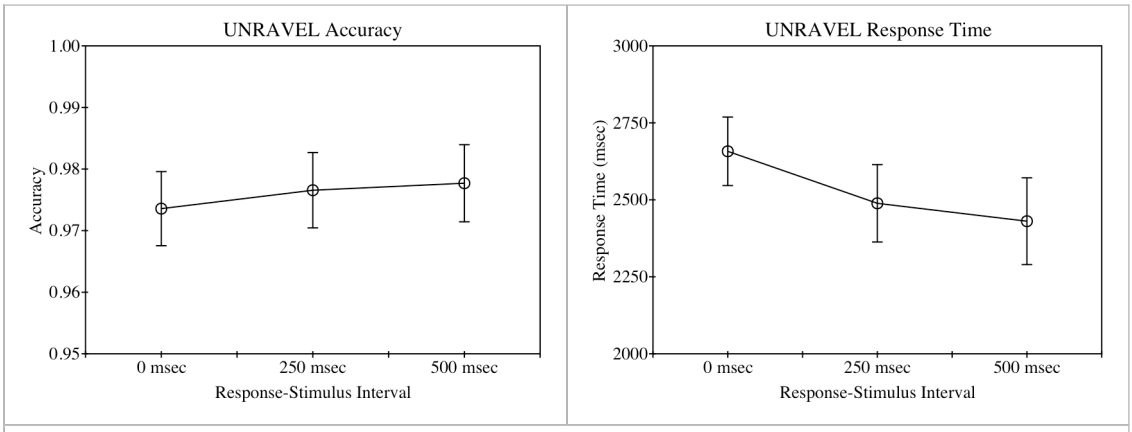


Figure 5: UNRAVEL accuracy (left) and response time (right), for trials not preceded by an interruption, as a function of response-stimulus interval in experiment 2. Error bars indicate one standard error. (cf. Figure 2.)

3.2 Results

All analyses proceeded as per experiment 1.

3.2.1 UNRAVEL Accuracy and Response Time

As for experiment 1, UNRAVEL trials with response time of greater than 20 seconds (less than 0.1% of trials) were excluded from analysis, as were those immediately following an interruption. On the remaining trials, accuracy was generally high (see figure 5, left), resulting in a negatively skewed distribution. As with experiment 1, a one-way within-subjects ANOVA on the accuracy data (arcsine transformed, to reduce skew) revealed no significant effect of response-stimulus interval ($F(2, 78) = 1.594$, $p = 0.210$, $\eta_p^2 = 0.039$). Also in agreement with experiment 1, response time decreased with response-stimulus interval, from a mean of 2658 msec to a mean of 2431 msec as response-stimulus interval increased from 0 msec to 500 msec (see figure 5, right), and a one-way within-subjects ANOVA revealed that the effect of response-stimulus interval on response time was significant ($F(2, 78) = 10.361$, $p < 0.001$, $\eta_p^2 = 0.210$). Follow-up t-tests indicated that response time was significantly slower when response-stimulus interval was 0 msec than when it was 250 msec ($t(39) = 3.847$, $p < 0.001$, two-tailed) or 500 msec ($t(39) = 4.113$, $p < 0.001$, two-tailed). The difference between response time when the response-stimulus interval was 250 msec compared to when it was 500 msec did not reach significance ($t(39) = 1.048$, $p = 0.301$, two-tailed).

A one-within ANOVA on the RT on correct baseline and post-interruption trials from blocks with RSI of 0 msec, revealed a significant effect of interruption length ($F(3, 117) = 3.997$, $p = 0.009$, $\eta_p^2 = 0.093$). As in experiment 1, the rank order of mean response times reflected interruption length (2658 msec, 2862 msec, 3163 msec, 3315 msec for trials following no, short, medium and long interruptions). Mean RT following long interruptions was significantly longer than that following short interruptions ($t(39) = 2.117$, $p = 0.041$, two-tailed) and baseline trials ($t(39) = 3.803$, $p < 0.001$, two-tailed).

While mean post-interruption response times were again slower for short response-stimulus intervals than for long response-stimulus intervals (3113 msec at 0 msec RSI versus 2959 msec at 500 msec RSI), a two-way within-subjects ANOVA revealed that the effect was not significant (main effect of response-stimulus interval: $F(2, 78) = 1.126$, $p = 0.329$, $\eta_p^2 = 0.028$). However, there was (as would be anticipated from the

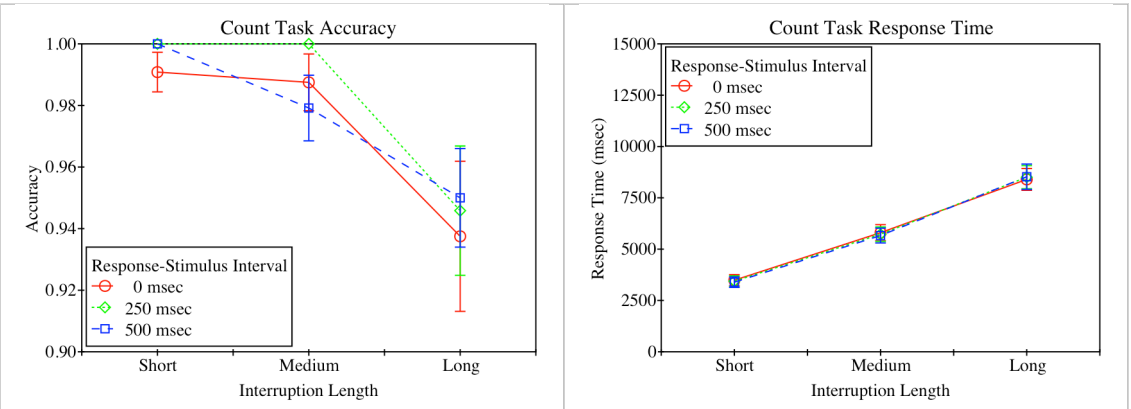


Figure 6: Interruption task accuracy (left) and response time (right) as a function of response-stimulus interval and interruption length in experiment 2. Error bars indicate one standard error. (cf. Figure 3.)

preceding analysis) an effect of interruption length ($F(2, 78) = 10.710$, $p < 0.001$, $\eta_p^2 = 0.215$) on post-interruption response time, with the longest interruptions leading to significantly longer response times than the shortest interruptions (3341 msec versus 2795 msec; $t(39) = 4.160$, $p < 0.001$). Response-stimulus interval and interruption length did not interact in their effects on post-interruption response time ($F(3.238, 126.267) = 0.731$, $p = 0.545$, $\eta_p^2 = 0.018$).

3.2.2 The Dot-Counting Task

Responses of more than 4 seconds per dot were excluded from analysis of performance on the dot-counting task (a total of 4 responses, all from the same participant). Accuracy (i.e., the proportion of dot-counting trials where the participant's answer was correct) and response times as a function of response-stimulus interval and interruption length are shown in figure 6. Interruption length affected accuracy ($F(1.174, 45.777) = 12.300$, $p < 0.001$, $\eta_p^2 = 0.240$) and response time ($F(2, 78) = 147.203$, $p < 0.001$, $\eta_p^2 = 0.791$), but response-stimulus interval did not (accuracy: $F(2, 78) = 0.904$, $p = 0.409$, $\eta_p^2 = 0.023$; response time: $F(2, 78) = 0.005$, $p = 0.995$, $\eta_p^2 = 0.000$). The variables also did not interact (accuracy: $F(2.764, 107.778) = 0.517$, p = 0.656, $\eta_p^2 = 0.013$; response time: $F(4, 156) = 0.649$, $p = 0.629$, $\eta_p^2 = 0.016$).[6]

Planned pair-wise comparisons confirmed that counting 12±1 dots took longer (mean: 8516 msec) than counting 8±1 dots (mean: 5774 msec; $t(39) \geq 8.099$, $p < 0.001$, in all cases), which took longer than counting 4±1 dots (mean: 3443 msec; $t(39) \geq 10.642$, $p < 0.001$, in all cases). Accuracy tended to be lower for longer interruptions (means: 94.4% for 12±1 dots, 98.9% for 8±1 dots, 99.7% for 4±1 dots), with accuracy on the longest interruptions being significantly lower than on the shortest interruptions, regardless of response-stimulus interval ($t(39) \geq 2.287$ $p \leq 0.028$, in all cases).

As in experiment 1, the analysis of response time confirms that the manipulation of the number of dots to be counted was effective in yielding temporally longer interruptions, though unlike experiment 1, accuracy on longer interruptions tended to be lower than accuracy on shorter interruptions. This is not surprising given the nature of the

[6] Accuracy data in this analysis were arcsine transformed to reduce skew.

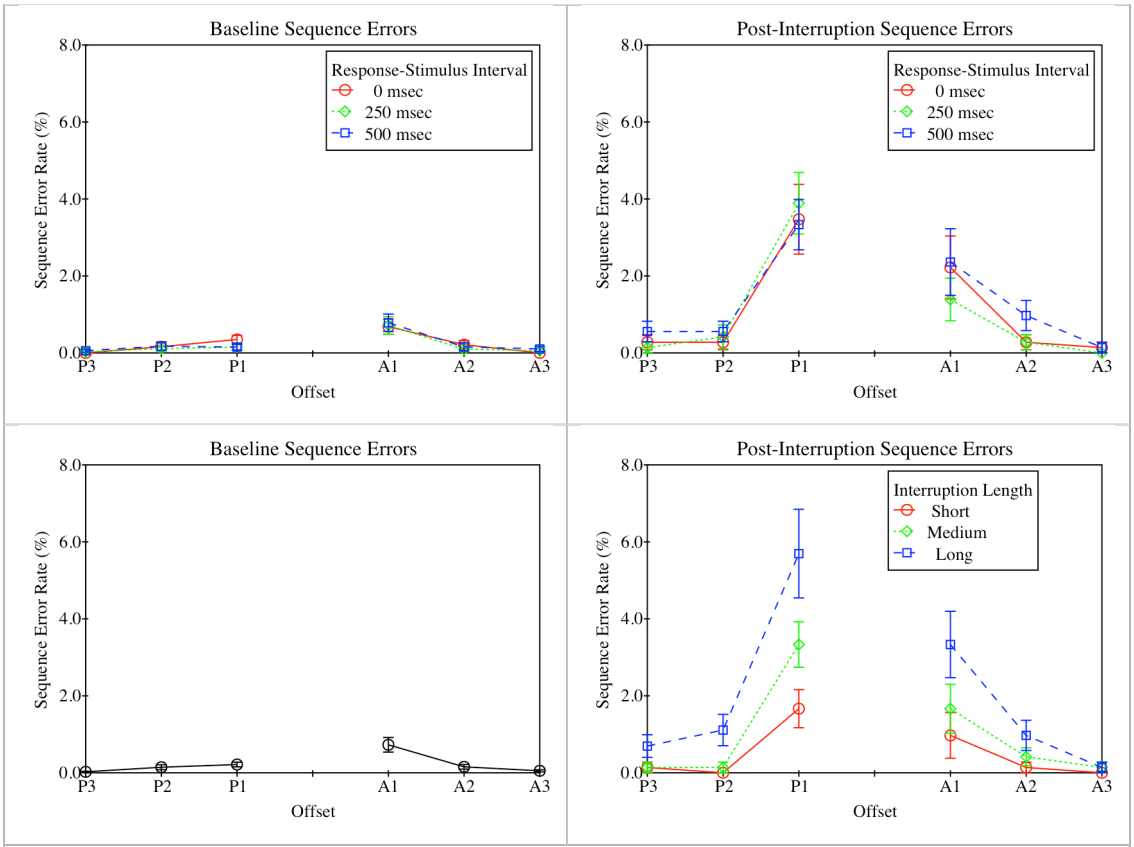


Figure 7: Sequence error rates on baseline UNRAVEL trials (i.e., those not preceded by an interruption) (left) and on post-interruption UNRAVEL trials (right) in experiment 2. The upper panels show the data by response-stimulus interval (pooled over interruption duration in the upper right panel), while the lower panels show the data pooled over response-stimulus interval. Error bars indicate one standard error. (cf. Figure 4.)

interrupting task – participants were more likely to err when counting approximately 12 dots than when counting approximately 4 dots.

3.2.3 Sequence Errors on the UNRAVEL Task

Figure 7 shows the proportion of errors of each type for baseline (i.e., non-interrupted) UNRAVEL trials (left) and post-interruption UNRAVEL trials (right, showing each level of interruption length) in experiment 2. As in the previous experiment, the number of opportunities to err in each cell of the experimental design (particularly following an interruption) was low. Therefore in both cases the data are pooled over response-stimulus interval for clarity. Echoing the analysis of sequence errors for experiment 1 and comparing P1 and A1 error rates (log transformed) as a function of the presence and length of preceding interruption, a 2 (error type) by 4 (interruption level) ANOVA revealed significant main effects of error type ($F(1, 39) = 10.657$, $p = 0.002$, $\eta_p^2 = 0.215$) and interruption level ($F(3, 117) = 17.038$, $p < 0.001$, $\eta_p^2 = 0.304$), as well as a significant interaction ($F(2.054, 80.119) = 3.692$, $p = 0.028$, $\eta_p^2 = 0.086$). Follow-up related t-tests comparing P1 and A1 error rates revealed a significant difference in favor of A1 errors on baseline trials ($t(39) = -3.450$, $p = 0.001$, two-tailed), a significant difference in favor of P1 errors following short interruptions ($t(39) = 2.090$, $p = 0.043$,

	P3	P2	P1	A1	A2	A3
RSI	0.046	0.021	0.019	0.035	0.166	0.026
IL	0.267	6.708	2.3×10^{6}	9.269	0.400	0.011
RSI × IL	4.4×10^{-4}	1.5×10^{-4}	0.002	0.004	0.004	1.3×10^{-4}

Table 2: Bayes Factors (BF_{Inc}) for inclusion of RSI, IL and their interaction in models of each sequence error type (experiment 2). Note that in situations where errors occur with sufficient frequency (P2, P1 and A1), the data provide strong evidence for models ($BF_{Inc} >> 1$) with a single factor (Interruption Length). In contrast, the results provide evidence against models ($BF_{Inc} << 1$) that include RSI or the RSI × IL interaction term.

two-tailed) and medium interruptions ($t(39) = 3.017$, $p = 0.004$, two-tailed) but no significant difference between error types following long interruptions ($t(39) = 1.577$, $p = 0.123$, two-tailed).

As in experiment 1, the second issue of concern relates to whether response-stimulus interval affects the rate of each type of sequence error, either on baseline trials or following an interruption. A 3 (response-stimulus interval) × 4 (interruption level: none, short, medium, long) multivariate analyses of variance (MANOVA) was performed on the (log transformed) proportions of each type of error (i.e., with six dependent measures: A1 errors, A2 errors, A3 errors, P1 errors, P2 errors and P3 errors). As was the case in experiment 1, this analysis revealed no significant interaction between response-stimulus interval and interruption length (Pillai's Trace = 0.730, $F(29, 11) = 1.026$, $p = 0.510$, $\eta_p^2 = 0.730$) and no significant effect of response-stimulus interval (Pillai's Trace = 0.291, $F(12, 28) = 0.959$, $p = 0.508$, $\eta_p^2 = 0.291$). In contrast to the absence of effects related to response-stimulus interval, the effect of interruption length was significant (Pillai's Trace = 0.673, $F(18, 22) = 2.518$, $p = 0.021$, $\eta_p^2 = 0.673$).

As with experiment 1, the results of this MANOVA must be interpreted with caution since the error data are sparse and skewed.[7] Again, the critical findings relate to the lack of an effect of response-stimulus interval in the presence of an effect of interruption length. Analogous Bayesian analyses to those following experiment 1 support the null effect of RSI or its interaction with interruption length (see table 2), at least for those errors that occur with sufficient frequencies to allow interpretation of the results (P2, P1, and A1). Moreover, the fact that the critical results echo those of experiment 1 further raises confidence in their reliability.

3.3 Discussion

While there are some differences between the results of experiment 1 and experiment 2, the gross patterns are the same: a) response-stimulus interval affects RT measures on the UNRAVEL task but not sequence errors, while b) interruption length affects post-interruption sequence errors, but c) the two manipulations do not interact in any analysis on any dependent measure. These findings support the view that the independent

[7] As in the case of experiment 1, reanalysis of the data excluding the no-interruption condition, i.e., via a 3 × 3 MANOVAs, reduces the violation of homogeneity of variance but yields qualitatively equivalent results.

variables affect distinct aspects of task-set processing that operate independently in the control of behavior on the interrupted UNRAVEL task.

The most obvious difference between the results of experiment 1 and experiment 2 relates to performance on the interrupting task. Participants made more errors when counting dots than when copying letters, and accuracy on dot-counting was particularly poor when the number of dots was high. Notwithstanding this, the manipulation of the number of dots achieved its purpose – more dots resulted in longer interruptions.

Turning to the interruption data, it is clear that the mean interruption length was somewhat shorter in experiment 2 (mean of 3443 msec, 5734 msec and 8516 msec for short, medium and long interruptions) than in experiment 1 (4895 msec, 8713 msec, and 12268 msec, respectively). However, this difference in duration is not clearly mirrored in sequence error rates (e.g., 0.97%, 1.67% and 3.33% for A1 errors in experiment 2 versus 0.83%, 1.67% and 4.03% for the equivalent in experiment 1). It appears that the interrupting task of experiment 2 produces slightly greater interference (per unit time of interruption) than that of experiment 1. This may be because the interrupting task in this experiment – dot counting – requires more cognitive processing than that of experiment 1 – copying characters. Alternatively (or in addition), it may be because dot counting potentially has a sequential element (at least when there are many dots and the task cannot be completed by subitizing), and it may be that this places demands on place-keeping, thus leading to greater interference.

4 Experiment 3

Experiments 1 and 2 show that the variables that affect place-keeping do not appear to affect set-shifting, and vice versa. They further suggest that these results are independent of the interrupting task. That is, not only are the processes supporting place-keeping and set-shifting independent or separable, but the nature of the interrupting task does not modulate the effects of either. Experiment 3 was designed to further explore this apparent independence of place-keeping and set-shifting by using a further interrupting task – in this case, a simple arithmetic task derived from that used in the study of interruptions by Botvinick and Bylsma (2005). In that study, participants repeatedly performed a relatively complex hierarchical sequential task (preparing instant coffee), with interruptions on 80% of trials. Interruptions were subtraction problems in which participants were required to calculate the change following a cash transaction (e.g., \$5.00 tendered for a purchase of \$1.60 → change is \$3.40). In order to control the length of the interruption experiment 3 used simple 1 digit addition and subtraction, but varied the number of problems (from 1 to 3) in order to vary the interruption length. Use of the arithmetic interrupting task is further motivated by the question of whether interrupting tasks with no obvious sequential character (in contrast to the task of experiment 2) would lead to similar results.

4.1 Method

4.1.1 Participants

Forty adult participants were recruited from Birkbeck's volunteer panel (27 female, mean age 32.5 years; 13 male, mean age 28.8 years). Participants were paid £7.50 to complete the study and all participants' data were included in the analyses.

4.1.2 Design

The design was the same as for experiments 1 and 2.

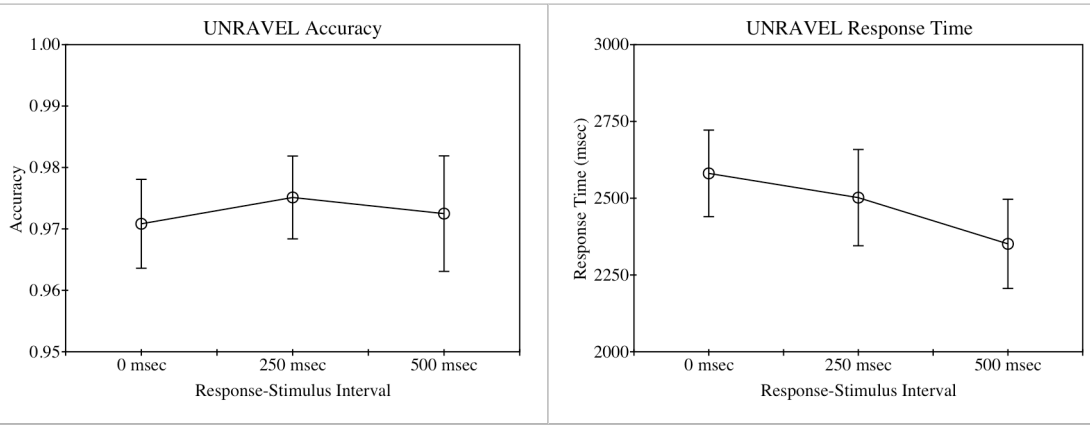


Figure 8: UNRAVEL accuracy (left) and response time (right), for trials not preceded by an interruption, as a function of response-stimulus interval in experiment 3. Error bars indicate one standard error. (cf. Figures 2 and 5.)

4.1.3 Interrupting Task: Single Digit Arithmetic

In experiment 3 the interrupting task consisted of one, two or three single digit addition or subtraction problems (e.g., 7 – 3 = ?) presented in succession. The problems appeared on screen and participants were required to type their response using the standard keyboard (and whichever fingers they chose) and then press the return key. As in the earlier experiments, participants were required to provide a response but that response was not required to be correct in order to proceed. Interrupting task accuracy was calculated as the proportion of arithmetic problems answered correctly within each interruption (i.e., 0/3, 1/3, 2/3, or 3/3).

4.1.4 Procedure

The procedure was identical to experiments 1 and 2 except that the interrupting task was replaced with the single-digit arithmetic task described above. On short interruption trials participants were asked to complete one single-digit arithmetic problem. On medium interruption trials two problems were presented in succession. On long interruption trials three problems were presented in succession. Participants were not informed of the number of arithmetic problems in any interruption – they were instructed to complete the arithmetic problems until an UNRAVEL stimulus reappeared. Thus, in contrast to experiments 1 and 2, pressing the return key did not always result in presentation of the next UNRAVEL stimulus.

As in the previous experiments, participants were given practice on the interrupting task prior to the main experiment (again as part of the second block of practice trials), with feedback given on the practice trials but not on the experimental trials. Also as with previous experiments, prior to embarking on data collection the experiment received ethical approval from the departmental ethics committee.

4.2 Results

4.2.1 UNRAVEL Accuracy and Response Time

As with the previous experiments, UNRAVEL trials with response time of greater than 20 seconds (less than 0.2% of trials) and those immediately preceded by an interruption were excluded from all analysis. On the remaining trials, accuracy was again generally high (see figure 8, left), resulting in a negatively skewed distribution. As in experiment 1, a one-way within-subjects ANOVA on the accuracy data (arcsine transformed, to

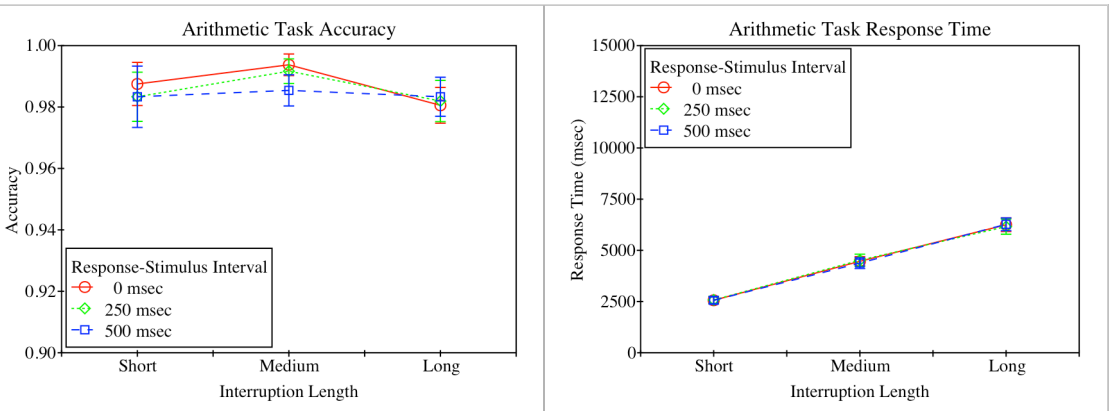


Figure 9: Interruption task accuracy (left) and response time (right) as a function of response-stimulus interval and interruption length in experiment 3. Error bars indicate one standard error. (cf. Figures 3 and 6.)

reduce skew) revealed no significant effect of response-stimulus interval ($F(2, 78) = 1.369$, $p = 0.260$. $\eta_p^2 = 0.034$). In agreement with both experiment 1 and experiment 2, response time decreased with response-stimulus interval, from a mean of 2581 msec to a mean of 2351 msec as response-stimulus interval increased from 0 msec to 500 msec (see figure 5, right), and a one-way within-subjects ANOVA confirmed that the effect of response-stimulus interval on response time was significant ($F(2, 78) = 7.879$, $p = 0.001$, $\eta_p^2 = 0.168$). Follow-up t-tests indicated that response time was significantly slower when the response-stimulus interval was 0 msec than when it was 500 msec ($t(39) = 4.292$, $p < 0.001$, two-tailed), as was the response time when the response-stimulus interval was 250 msec compared to 500 msec ($t(39) = 2.451$, $p = 0.019$, two-tailed). The difference between response time when the response-stimulus interval was 0 msec compared to when it was 250 msec did not reach significance ($t(39) = 1.295$, $p = 0.203$, two-tailed).

A one-within ANOVA on the RT on correct baseline and post-interruption trials from blocks with RSI of 0 msec, revealed a significant effect of interruption length ($F(3, 117) = 12.834$, $p < 0.001$, $\eta_p^2 = 0.248$). Mean RT on the UNRAVEL task was slower in trials preceded by a short, medium or long interruption (3215 msec, 3157 msec, 3239 msec) than in baseline trials (2581 msec; all $t(39) > 4.917$, $p < 0.001$, two-tailed), but interruption length did not affect RT.

Mean post-interruption response times were again slower for short response-stimulus intervals than for long response-stimulus intervals (3084 msec at 0 msec RSI versus 3211 msec at 500 msec RSI), but a two-way within-subjects ANOVA revealed that the effect was not significant (main effect of response-stimulus interval: $F(1.703, 66.435) = 1.421$, $p = 0.248$, $\eta_p^2 = 0.035$). The effect of interruption length on post-interruption response times was also not significant ($F(2, 78) = 1.543$, $p = 0.220$, $\eta_p^2 = 0.038$), nor was the interaction between these variables ($F(4, 156) = 1.621$, $p = 0.172$, $\eta_p^2 = 0.040$).

4.2.2 The Arithmetic Task

Responses of more than 10 seconds per arithmetic problem were excluded from analysis of performance on the arithmetic task (a total of 6 responses, 5 of which were from a single participant). Accuracy and response times as a function of response-stimulus

interval and interruption length are shown in figure 9. Interruption length had a marginally significant effect on accuracy[8] ($F(2, 78) = 2.981$, $p = 0.057$, $\eta_p^2 = 0.071$) and a significant effect on response time ($F(1.408, 54.906) = 317.531$, $p < 0.001$, $\eta_p^2 = 0.891$). Response-stimulus interval did not affect either accuracy ($F(2, 78) = 0.126$, $p = 882$, $\eta_p^2 = 0.003$) or response time ($F(2, 78) = 0.048$, $p = 0.953$, $\eta_p^2 = 0.001$). The independent variables also did not interact on either measure (accuracy: $F(4, 156) = 0.535$, $p = 0.710$, $\eta_p^2 = 0.014$; response time: $F(2.978, 116.158) = 0.829$, $p = 0.480$, $\eta_p^2 = 0.021$).

Planned comparisons confirmed that completing 3 arithmetic problems took longer (mean: 6213 msec) than completing two (mean: 4452 msec; $t(39) \geq 9.684$, $p < 0.001$ in all cases), which took longer than completing one (mean: 2560 msec; $t(39) \geq 11.456$, $p < 0.001$, in all cases). As in the case of the earlier experiments, the analysis of response time confirms that the manipulation of the number of arithmetic problems was effective in yielding longer interruptions. Again, this is to be expected given the nature of the interrupting task.

4.2.3 Sequence Errors on the UNRAVEL Task

Figure 10 shows the proportion of errors of each type for baseline (i.e., non-interrupted) UNRAVEL trials (left) and post-interruption UNRAVEL trials (right, showing each level of interruption length) in experiment 3. As in figures 4 and 7, the upper panels show the (lack of) effect of response-stimulus interval while the data in the lower panels are pooled over response-stimulus interval for clarity. Echoing the analysis of sequence errors for the previous experiments, comparing P1 and A1 error rates (log transformed) as a function of the presence and length of preceding interruption, a 2 (error type) by 4 (interruption level) ANOVA revealed no significant effect of error type ($F(1, 39) = 0.028$, $p = 0.867$, $\eta_p^2 = 0.001$), a main effect of interruption level ($F(2.011, 78.441) = 7.541$, $p < 0.001$, $\eta_p^2 = 0.162$), and no significant interaction ($F(2.193, 85.518) = 1.808$, $p = 0.150$, $\eta_p^2 = 0.044$). Follow up related t-tests comparing P1 and A1 error rates (not licensed by the ANOVA, but reported for completeness and comparison with the previous studies) revealed a significant difference in favor of A1 errors on baseline trials ($t(39) = -2.663$, $p = 0.011$, two-tailed), but no significant difference following short interruptions ($t(39) = -0.954$, $p = 0.346$, two-tailed), medium interruptions ($t(39) = 0.784$, $p = 0.438$, two-tailed), or long interruptions ($t(39) = 1.409$, $p = 0.167$, two-tailed).

As in the previous experiments, a multivariate ANOVA with (log transformed) error type (P3, P2, P1, A1, A2, A3) as the dependent measures and response-stimulus interval (3 levels) and interruption length (4 levels: none, short, medium, long) as within-subjects factors was conducted to determine whether interruption length affected the proportion of any error type (either singly or in conjunction with response-stimulus interval). There were no significant effects when all six dependent measures were entered into the MANOVA (interruption length: Pillai's Trace = 0.513; $F(16, 24) = 1.581$, $p = 0.151$, $\eta_p^2 = 0.513$; response-stimulus interval: Pillai's Trace = 0.304; $F(10, 30) = 1.313$, $p = 0.268$. $\eta_p^2 = 0.304$; the interaction between response-stimulus interval and interruption length: Pillai's Trace = 0.561; $F(23, 17) = 0.944$, $p = 0.559$, $\eta_p^2 = 0.561$). Note however that the durations of interruptions in this experiment were lower than in experiments 1 and 2. Post-interruption error rates were correspondingly lower, with few errors occurring at offsets of greater than 1. A MANOVA restricted to just P1

[8] Accuracy data were again arcsine transformed prior to analysis to reduce skew.

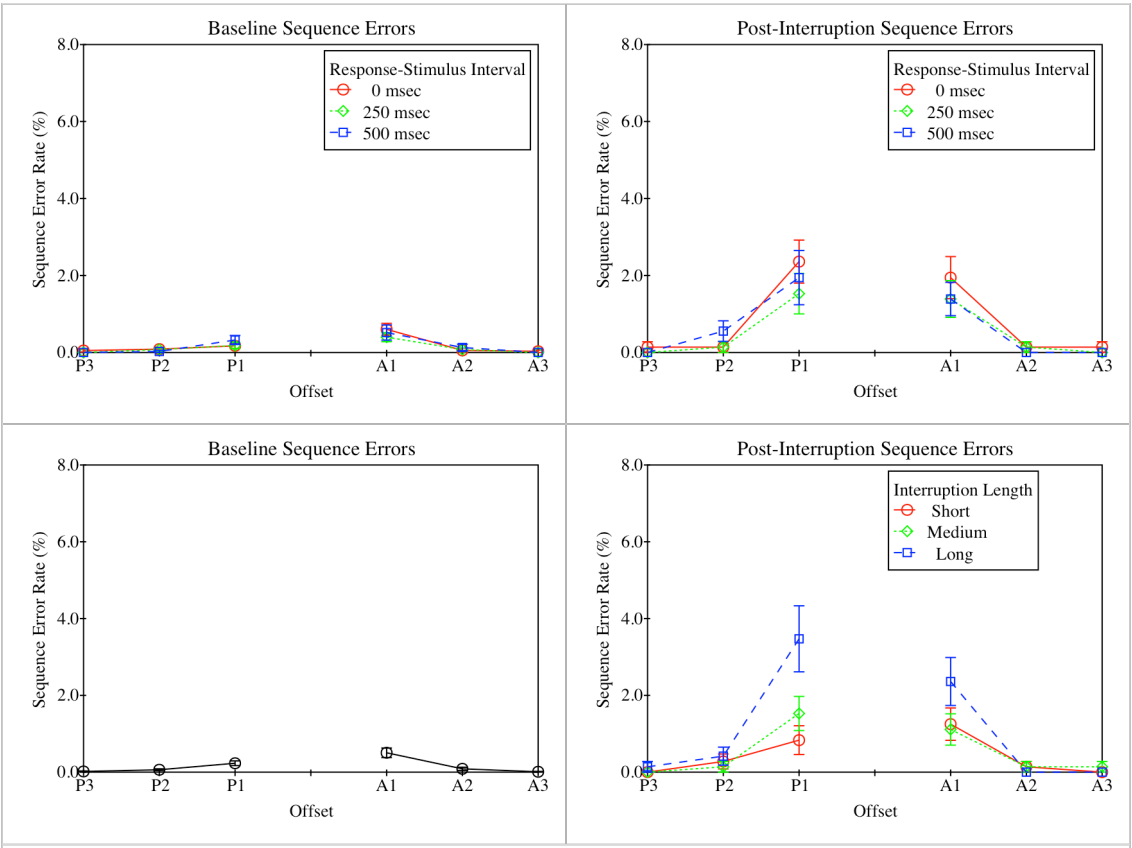


Figure 10: Sequence error rates on baseline UNRAVEL trials (i.e., those not preceded by an interruption) (left) and on post-interruption UNRAVEL trials (right) in experiment 3. The upper panels show the data by response-stimulus interval (pooled over interruption duration in the upper right panel), while the lower panels show the data pooled over response-stimulus interval. Error bars indicate one standard error. (cf. Figures 4 and 7.)

and A1 error types yielded a significant effect of interruption length (Pillai's Trace = 0.347; $F(6, 34) = 3.015$, $p = 0.018$, $\eta_p^2 = 0.347$) but no significant effect of response-stimulus interval (Pillai's Trace = 0.158; $F(4, 36) = 1.685$, $p = 0.175$, $\eta_p^2 = 0.158$) and no interaction between the factors (Pillai's Trace = 0.267; $F(12, 28) = 0.849$, $p = 0.604$, $\eta_p^2 = 0.267$).

Following the Bayesian analyses of experiments 1 and 2, a series of Bayesian repeated measures analyses of variance were conducted using the JASP package (JASP Team, 2018; Version 0.8.5.1), with response-stimulus interval and interruption length as within-subjects factors. Bayes factors for inclusion of each term in the various models are shown in Table 3. The low error rates limits interpretation of the results, but for the two most frequent types of errors (P1 and A1) there is (respectively) extremely strong and anecdotal evidence for an effect of interruption length on error production in the presence of strong evidence against an effect of response-stimulus interval or an interaction between the factors. While these results are less strong than those of experiments 1 and 2, in that they do not relate to errors at each offset, they nevertheless further bolster the results of experiment 1 and experiment 2 concerning a) the absence of any relation between response-stimulus interval and sequence errors, and b) the

	P3	P2	P1	A1	A2	A3
RSI	0.052	0.060	0.030	0.028	0.022	0.048
IL	0.020	0.018	5665.833	1.735	0.010	0.021
RSI × IL	2.7×10^{-4}	7.1×10^{-4}	0.003	6.4×10^{-4}	1.1×10^{-4}	2.8×10^{-4}

Table 3: Bayes Factors (BF_{Inc}) for inclusion of RSI, IL and their interaction in models of each sequence error type (experiment 3). While the low rates of P3, P2, A2 and A3 errors limit interpretation, there is strong evidence for an effect of interruption length on P1 ($BF_{Inc} >> 1$) and anecdotal evidence for the effect on A1 errors. In contrast (and where interpretable), the results provide strong evidence against models ($BF_{Inc} << 1$) that include RSI or the RSI × IL interaction term.

tendency for longer interruptions to result in more one-step errors than shorter interruptions, and more P1 errors than A1 errors.

4.3 Discussion

The results of experiment 3 support the key claims made on the basis of experiments 1 and 2, namely that a) the effect of an interrupting task on sequence errors is to increase the probability of perseverative errors more than that of anticipatory errors, b) the effects of interruptions are quantitatively more pronounced but not qualitatively different for longer interruptions, and c) the processes that support set-shifting are primarily sensitive to RSI and not interruption length, while the processes that support place-keeping are sensitive to interruption length but not RSI.

The mean duration of interruptions in experiment 3 (2560 msec, 4452 msec and 6214 msec for short, medium and long interruptions, respectively) was less than that of the earlier experiments, but the effect on sequence errors was similar to that of experiment 2. Thus the longest interruptions in experiment 3 (of mean duration 6.214 seconds) led to a total sequence error rate of 6.4%, compared with a medium interruption in experiment 2 (of mean duration 5.774 seconds) that led to a total sequence error rate of 5.8%. Differences between the interrupting tasks of experiments 2 and 3 therefore do not appear to have substantially impacted sequence error rates. This suggests that the interference from dot-counting in experiment 2 is more related to the cognitive aspects of the dot-counting task, rather than to its possible involvement of place-keeping, though clearly additional data that directly speaks to this issue is needed to fully adjudicate between these possibilities.

A second distinctive feature of experiment 3 compared to the earlier experiments is that response time on post-interruption trials was not dependent on interruption length: all levels of interruption length led to a similar lengthening of response times of approximately 600 msec in comparison to baseline trials. One possible reason for the lack of an effect of interruption length in this experiment is that, in contrast to the earlier experiments, participants could not be sure when pressing the return key during an interruption that they would subsequently be presented with an UNRAVEL stimulus – frequently they would be presented with another arithmetic problem. Consequently this experiment may provide a more pure measure of the time cost of recovering from an interruption than that provided by the earlier experiments.

The fact that, in this experiment, pressing the return key on completion of an interrupting arithmetic problem did not predictably lead to resumption of the UNRAVEL task also speaks to another potential interpretation of participant performance. In experiments 1 and 2 it is possible that participants might, on completion of the interrupting task (typing the letters or entering the count of the number of dots) pause and attempt to recall their position in the UNRAVEL task before pressing the return key and resuming the UNRAVEL task. Our primary evidence for the position presented here is post-interruption sequence errors, rather than post-interruption response times, and so if participants were to adopt this strategy it would not undermine our theoretical position. However, the strategy is not effective in the current experiment. At the same time, all previous critical results hold, suggesting that, if such a strategy is used in the earlier experiments, it does not affect the critical findings.

5 Interim Discussion and Additional Analyses

5.1 Summary and Interpretation of Results

Across three experiments we have shown that while increasing response-stimulus interval (RSI) reduces response time on the UNRAVEL task, it does not affect the number or type of errors following an interruption. This was true for all levels and types of interruption investigated. At the same time, RSI does affect response time on the UNRAVEL task in precisely the way one would anticipate if performance of the UNRAVEL task relies on the same mechanisms underlying the performance of more standard set-shifting tasks. That is, UNRAVEL response times are faster at longer RSIs, but the decrease in response time with increasing RSI does not fully reflect the increase in RSI, meaning that longer RSIs appear to allow some, but not all, processes required for set-shifting (e.g., disengagement of the previous task-set or activation of the forthcoming task-set) to be carried out during the response-stimulus interval.

In addition, across all tasks, long interruptions led (unsurprisingly) to more post-interruption sequence errors, and replicating previous work, post-interruption sequence errors tended to be perseverative rather than anticipatory, with long interruptions resulting in disproportionately more perseverative than anticipatory errors. Critically, however, there was no evidence from any of the three studies that this finding is modulated by RSI. That is, while longer RSI reduces response time on the UNRAVEL task, it does not translate into fewer sequence errors on that task, either in the baseline condition or following an interruption. This contrasts with the view that longer RSIs might allow some degree of task-set configuration to occur prior to an interruption, which facilitates place-keeping, and thereby reduces the rate of affect post-interruption sequence errors. The implication is that, at some level, the processes supporting task-set reconfiguration and place-keeping operate on distinct representations of task-set.

One potential concern with the above interpretations is that the factorial design of each experiment meant that there were relatively few interruptions of each duration at each level of RSI, and hence relatively few opportunities for post-interruption error. Given that the findings were consistent across all three experiments, that in most cases the critical F ratios were less than one, and that our interpretation is supported by the Bayesian analyses, we consider it unlikely that the results are due to a lack of power, though it is possible that the manipulation of RSI is simply too weak to allow a detectable effect on sequence errors. Recall however that all three experiments took the

same form and that all three yielded directly comparable datasets. In order to increase statistical power and confidence in our interpretation we therefore conducted further analyses collapsing the data across all experiments (and using experiment as a between-subjects factor in the analyses).

For direct comparison with the earlier results, a 3 (experiment) × 3 (response-stimulus interval) × 4 (interruption level: none, short, medium, long) multivariate analysis of variance (MANOVA) was performed on the (log transformed) proportions of each type of error (i.e., with six dependent measures: A1 errors, A2 errors, A3 errors, P1 errors, P2 errors and P3 errors) across the three experiments. Consistent with previous results, this analysis revealed no significant interaction between response-stimulus interval and interruption length (Pillai's Trace = 0.252, $F(35, 83) = 0.798$, $p = 0.770$, $\eta_p^2 = 0.252$) and no significant effect of response-stimulus interval (Pillai's Trace = 0.075, $F(12, 106) = 0.716$, $p = 0.733$, $\eta_p^2 = 0.075$). In contrast to the absence of effects related to response-stimulus interval, the effect of interruption length was significant (Pillai's Trace = 0.555, $F(18, 100) = 6.932$, $p < 0.001$, $\eta_p^2 = 0.555$). In addition, within this analysis the three-way interaction between response-stimulus interval, interruption length and experiment was not significant (Pillai's Trace = 0.663, $F(70, 168) = 1.190$, $p = 0.184$, $\eta_p^2 = 0.331$), nor was the two-way interaction between response-stimulus interval and experiment (Pillai's Trace = 0.258, $F(24, 214) = 1.323$, $p = 0.151$, $\eta_p^2 = 0.129$), or the two-way interaction between interruption length and experiment (Pillai's Trace = 0.334, $F(36, 202) = 1.167$, $p = 0.250$, $\eta_p^2 = 0.172$). However, the main effect of experiment was significant (Pillai's Trace = 0.175, $F(12, 226) = 1.803$, $p = 0.049$, $\eta_p^2 = 0.087$). This latter effect reflects the fact that participants in experiment 1 made more errors than those in experiment 2, who made more errors than those in experiment 3, presumably because the interruption duration (for all three levels of interruption) was greater in experiment 1 than in experiment 2, and greater in experiment 2 than in experiment 3.

UNRAVEL response times as a function of interruption length also afford a cross-experiment analysis. In particular, in each of the preceding experiments we observed that UNRAVEL response time on post-interruption trials decreased with increasing interruption length, though in all three cases the effect was not significant. The increased power obtained by analyzing across experiments, however, suggests that the effect is real. Thus a 3 way ANOVA on post-interruption response time on the UNRAVEL task, with RSI (0 msec, 250 msec, 500 msec) and interruption duration (short, medium or long) as within-subjects factors and experiment (1, 2 and 3) as a between-subjects factor, revealed that not only is the main effect of interruption length significant ($F(1.877, 219.581) = 20.765$, $p < 0.001$, $\eta_p^2 = 0.151$), but that the main effect of RSI is also significant ($F(1.903, 222.608) = 4.156$, $p = 0.018$, $\eta_p^2 = 0.034$). The only other significant effect from this analysis relates to the interaction between interruption length and experiment ($F(3.754, 219.581) = 2.634$, $p = 0.038$, $\eta_p^2 = 0.043$). The significant effect of RSI on post-interruption response time is intriguing as it suggests that processing related to set-shifting *prior* to the interruption can carry over through the duration of the interruption. This effect argues against an account that attempts to explain our findings purely in terms of dissipation of task set throughout the duration of the response-stimulus interval and the interrupting task. The response-stimulus interval is generally small compared to the interruption length, and so the additional dissipation of task set due to variation of the RSI (by up to 500 msec) in addition to the interruption

interval (of over 12 seconds in the long-interruption condition of experiment 1) would be expected to be minimal, and would not be expected to affect post-interruption response time. Rather, the effect suggests some form of task-set preparation during the response-stimulus interval (prior to an interruption) which primes the upcoming task-set, with that priming persisting over the interruption interval. However, that priming (which would presumably be greater for longer RSI) does not modulate the post-interruption sequence error rate.

5.2 Implications for Information Processing Models of Set-Shifting and Place-Keeping

The implications of our results for the relationship between the cognitive processes that support set-shifting and place-keeping may be clarified by considering current theoretical accounts of those processes (or equivalently, the mechanisms held to give rise to set-shifting and place-keeping effects). Consider first set-shifting. While there are several information processing accounts of the effects of RSI on response time in set-shifting tasks, those accounts share some key features. Gilbert and Shallice (2002), for example, build on the influential interactive activation model of Stroop effects of Cohen et al. (1990). The Gilbert and Shallice model comprises input units for each value of each critical dimension of the stimulus, response units for each response, and task demand units which effectively amplify or prime the input units of the dimension (or dimensions) relevant to the current task. In the paradigm of interest, a task cue is first presented (indicating which task is to be performed) and subsequently a stimulus is presented. The interval between task cueing and stimulus presentation is manipulated. When this interval is non-zero, task demand units can accumulate activation prior to presentation of a stimulus. This leads to relatively fast response times when the stimulus is eventually presented, but the activation of task demand units is not instantaneous. Hence longer cue/stimulus intervals allow greater accumulation of task demand unit activation prior to stimulus presentation than shorter cue/stimulus intervals, and this is why, according to the model, increasing the cue/stimulus interval results in shorter response times once the stimulus is actually presented.

Altmann and Gray (2008) present an alternative model of set-shifting in which task control codes play a similar role to the task demand units of Cohen et al. (1990) and Gilbert and Shallice (2002). Like task demand units, task control codes are held to have associated activation values, but in the Altmann and Gray model task control codes are held in episodic memory, and a task control code's activation value determines whether it can be recalled. On stimulus presentation, the model attempts to recall the current task control code and then uses that code to generate an appropriate response. The recall process may need to be performed multiple times if it fails to produce a task control code. Increasing the cue/stimulus interval effectively allows task control codes to be activated prior to stimulus presentation, increasing the likelihood that recall of the current task control code will be effective. The account thus shares important features with that of De Jong (2000).

While there are theoretical differences between these two models of cue/stimulus presentation interval effects, they both propose that behavior is controlled by task-specific control units or codes, and that the activation of these units/codes increases during the interval between cue presentation and stimulus presentation. Within our version of the UNRAVEL task, this interval corresponds to the response-stimulus interval, as responding to one stimulus implicitly cues the next step of the UNRAVEL task sequence.

Building upon the model of Altmann and Gray (2008), Altmann and Trafton (2015) propose a two-stage model – the *remember-advance* model – of performance on the UNRAVEL task, and of place-keeping more generally. In this model each step of the UNRAVEL task has an associated control code. These are the same control codes postulated in the model of Altmann and Gray. On each step, the most active control code is recalled from episodic memory (as in the Altmann and Gray model). This is the "remember" stage of the remember-advance model. Typically the recalled control code will be the control code for the most recently completed step. That control code is then used to generate the control code for the current step (the "advance" stage), which is then executed, leaving a trace in episodic memory. Decay operates on the traces of control codes in episodic memory, so the most recent trace is normally the most active, but retrieval is also subject to noise, meaning that the *remember* stage may retrieve the trace of an earlier control code, rather than that of the most recent control code. This is held to be the root of perseverative errors. Generation of the control code for the current step relies on activation spreading from the retrieved control code to subsequent control codes – a process that is held to operate through directed associative links between successive control codes. So in the UNRAVEL task associative links will result in activation spreading from the U control code to the N control code, and from the N control code to the R control code, and so on. Activation spreads along chains of links, meaning that, for example, the U control code will activate the N control code which will activate the R control unit and so on, but the strength of the links is held to be less than 1, so if the U control code is the outcome of the *remember* stage, the N control code will be the most active at the *advance* stage, though noise in this process may result in the R control code (or even the A control code) being selected. This is held to be the cause of anticipation errors.

In support of their model, Altmann and Trafton (2015) report a study of interruptions in the UNRAVEL task where 1 second lags were added between steps, with probability of 0.5. Participants made fewer perseverative errors when interrupted following a lag than when interrupted on a no-lag trial. Altmann and Trafton suggest that these lags allow participants to effectively consolidate the most recently completed control code, thereby ensuring that such codes are more accurately recalled, and thereby reducing the rate of perseverative errors. The self-report of participants supports this account.

While the remember-advance model can provide an adequate account of several empirical effects, it lacks parsimony: perseverative and anticipatory errors are held to arise from separate mechanisms. Other models of sequential behavior in other domains (e.g., the model of serial recall of Burgess & Hitch, 1999) view both types of error as resulting from a single mechanism – a noisy representation of task or temporal context where nearby elements (both past and future) have similar (and hence confusable) representations. Similarity of these representations is held to result in the activation of the representations of nearby elements, with the level of activation depending on the similarity of task or temporal context. Within a model of UNRAVEL, such a mechanism would naturally result in error "gradients", with P1 errors more likely than P2 errors, and P2 errors more likely than P3 errors, and similarly with A1 errors more likely than A2 errors, and A2 errors more likely than A3 errors. However, as Altmann et al. (2014) point out, a single mechanism account does not sit well with asymmetries in perseverative and anticipatory errors, as observed both in their original study and as replicated here (cf. figures 4, 7 and 10), and indeed as observed in the study of lags by Altmann and Trafton (2015).

The data presented here, however, appear to present a difficulty for the remember-advance model, particularly if the task control codes of that model are identified with the task control codes of the model of Altmann and Gray (2008). Note that both models assume such codes are stored in episodic memory, that they have associated activation values, and that successful task performance requires recall of an appropriate task control code. The difficulty is that this approach would imply that increasing the RSI in the UNRAVEL task should result in greater activation of task control codes (both in general and prior to an interruption), and hence greater resistance to perseverative sequence errors, both during baseline performance and following an interruption. Our data instead suggest a separation between task control codes and the units responsible for priming responses when RSI is varied, with the activation of the units responsible for priming responses being inaccessible to the mechanisms that activate task control codes.

A further issue, somewhat orthogonal to the above, concerns the main effect of RSI on post-interruption RT found in the final cross-experiment analysis. This is worthy of special consideration given that post-interruption trials were not immediately preceded by a response-stimulus interval. (Recall that the interval occurred after the preceding UNRAVEL trial and before the interruption, but not after the interruption.) This is the only analysis that supports an effect of RSI that persists throughout the interruption period. It appears that processing after the preceding UNRAVEL trial but before the interruption leaves the cognitive system in a state of heightened readiness for the post-interruption UNRAVEL trial. Critically, however, this does not translate into a reduction in post-interruption errors. In other words, this analysis also supports the interpretation that place-keeping and set-shifting effects arise from separate task-set representations. The task-set representation that persists during the interruption period does not appear to be accessible to the mechanisms that support place-keeping.

A final puzzle concerning these models and datasets is that our results (i.e., that there is no effect of increasing RSI on sequence error rates) might seem to contradict the results of Altmann and Trafton (2015), where (as noted) adding lags before some UNRAVEL trials led to a reduction in perseverative errors when a lag was followed by an interruption. One possibility is that the irregular nature of Altmann and Trafton's task encouraged the use of a rehearsal or consolidation strategy on lag trials (as reported by their participants), while this strategy was not encouraged by the more regular / rhythmic nature of our procedure. Another possibility is that our maximum response-stimulus interval (of 500 msec) was simply too short to make this deliberate strategy viable.

5.3 An Alternative Account

Our data, together with past results, suggest some critical features required of an integrated model of set-shifting and place-keeping. First, the data suggest separable codes for tracking one's position in a task and for controlling task-set engagement when shifting from one set to another. Altmann and Trafton's (2015) work suggests that the former are open to deliberate rehearsal, and so one possibility is that these are maintained within a high-level deliberative control system (e.g., a task-general working memory, or possibly the supervisory system, of Norman & Shallice, 1986). Equally there is reason to believe that the latter are not accessible to deliberate processes, and within the dual-systems architecture of Norman and Shallice (1986) these would presumably be located within the contention scheduling system (i.e., a lower-level system responsibly for the execution of routine behavior, potentially but not necessarily

under indirect control of the supervisory system). Critically, activation of representations in the contention scheduling system is not directly accessible to the supervisory system.

Second, the gradient over error rates, described in the previous section, suggests that place-keeping may rely on a noisy representation of task or temporal context where nearby steps have similar representations, as employed, for example, in the memory model of Burgess and Hitch (1999).[9]. The asymmetry in error types, we suggest, arises not through distinct mechanisms for retrieval of episodic task control codes and advancing to the next step, but through this noisy representation of task context, combined with a (possibly deliberate) mechanism for "ticking off" or inhibiting recently completed steps. Interruptions, we suggest, disrupt both the representation of temporal context, leading to a general increase in sequence errors, and this inhibitory mechanism (which normally prevents perseverative errors), leading to an increase in perseverative errors.

Finally, we assume that the place-keeping mechanism is event-based rather than time-based, such that updating of the context signal is dependent upon the occurrence of events (e.g., generation of a response) rather than time (cf. Koch et al., 2010). In contrast, activation-based processing at the lower level is, we assume, time-based (following Gilbert & Shallice, 2002, and related models). These assumptions mean that place-keeping will be adversely affected by intervening tasks but not by a lengthening of the response-stimulus interval. In contrast set-shifting will be facilitated by increasing the response-stimulus interval, which effectively allows disengagement of previous task sets and/or, in a predictable task, priming or preparation of the upcoming task-set.

6 A Model of Place-Keeping and Set-Shifting in the UNRAVEL Task

In order to demonstrate the viability of the above alternative account, and to specify it more fully, this section presents a detailed computational implementation of the account, together with simulations of the key findings from the empirical work reported above. We take a step-wise approach, beginning by extending the model of task switching of Gilbert and Shallice (2002) to the seven UNRAVEL subtasks, and assuming perfect place-keeping but no interruptions. We then extend the model with our proposed place-keeping mechanism before considering how this mechanism (and the lower-level mechanisms that are hypothesized to give rise to set-shifting effects) is affected by interruptions of varying length.

6.1 Extending the Gilbert and Shallice (2002) Model to the UNRAVEL Task

As discussed above, Gilbert and Shallice (2002) present a model of switching between two tasks (color naming and word reading in Stroop). The model shown in figure 11, which we refer to as Model A, presents a direct extension of this model to the UNRAVEL task. Model A consists of three banks of units: 14 input units (organized as 7 pairs of units, corresponding to the 7 dimensions of an UNRAVEL stimulus); 14

[9] This is also consistent with supplementary analyses of the results from the current experiments, where in all cases strong positive correlations were observed between the rates of post-interruption anticipatory and perseverative errors (r = 0.497, 0.306, 0.562, and p = 0.001, 0.055, 0.001, for experiments 1, 2 and 3 respectively, with N = 40 in all cases). Such correlations are suggestive of a common mechanism behind both types of error.

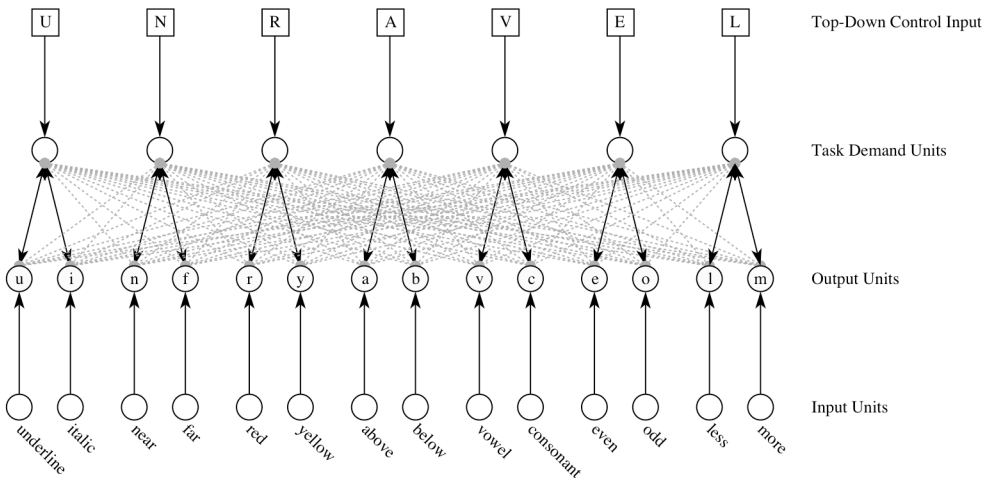


Figure 11: Model A: A model of the subtasks of UNRAVEL, based on the task switching model of Gilbert and Shallice (2002).

output units, corresponding to the 14 possible responses in the UNRAVEL task and again organized as 7 pairs of units, with each pair corresponding to the one UNRAVEL subtasks; and 7 task demand units (corresponding to the 7 UNRAVEL subtasks). Each unit has an activation value that varies over time, as in the original model, between –1 and +1, as a result of excitation and inhibition from connected units. Input units are connected to corresponding output units, and output units are bidirectionally connected to corresponding task demand units. In addition, lateral inhibition operates within each pair of output units and between task demand units. Task demand units also inhibit output units corresponding to other tasks. These connections are reciprocal, with output units associated with each task inhibiting the task demand units for the other tasks.

Model A differs from the model of Gilbert and Shallice (2002) in that a) there are seven tasks instead of two, b) each task has two output units, corresponding to each binary stimulus dimension, and c) trainable connections between input units and task demand units are not included. With respect to the last of these differences, Gilbert and Shallice (2002) include trainable connections from input units to task demand units in their model in order to model item-specific effects. These connections have been excluded because such effects are not the focus of the current discussion. In all other respects, including the calculation of activation and the propagation of activation throughout a trial, Model A is a direct translation of the Gilbert and Shallice (2002) model to the UNRAVEL task. (See the appendix for full details.)

Model A requires specification of task-specific top-down control input strengths and task-specific input to output connection strengths. In order to achieve modest rates of task errors (e.g., responding ‘U’ on a U/I discrimination subtask when the correct response was ‘I’) similar to those of participants on the UNRAVEL task (i.e., around 1%), the former were set to random values between 14.0 and 16.0 while the latter were set to values between 0.9 and 1.0. With these task-specific parameter values, and lateral inhibition set to 4.0 (in contrast to 2.0 of the original model), the bias on task-demand units set to –10 and the bias on output units set to –9 (in contrast to –4 and –6, respectively, in the original model, to reflect the increased number of task demand

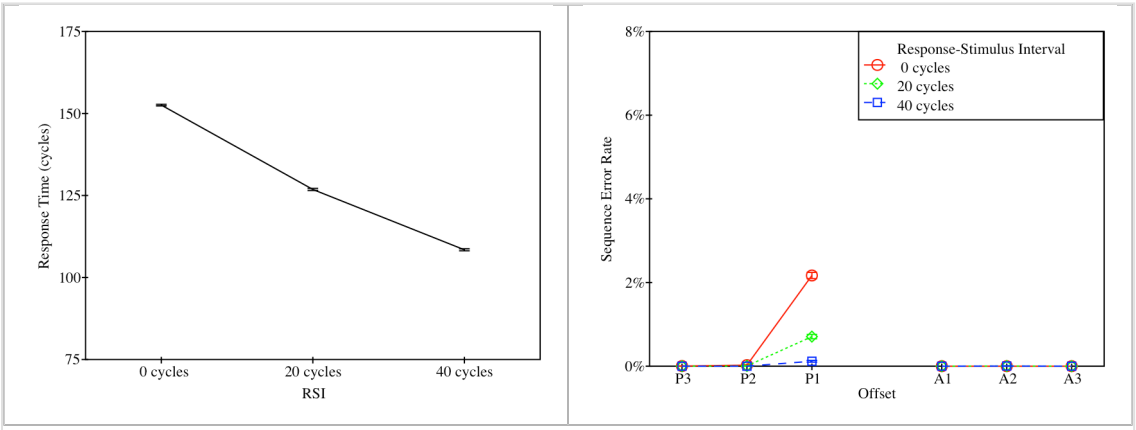


Figure 12: Predictions of Model A. Left panel: Effect of RSI on response time (measured in processing cycles). Right panel: Effect of RSI on the percentage of different types of sequence error in uninterrupted trials of the UNRAVEL task. The data are based on 40 virtual participants, with parameters set as in Gilbert and Shallice (2002) except where explicitly mentioned in the text. Sequential subtask selection was controlled by an error-free mechanism that took 50 processing cycles on each trial to select the appropriate subtask. Error-bars represent one standard error about the mean.

units[10]), the model is able to perform each of the UNRAVEL sub-tasks. Moreover, if the top-down control units are activated in sequence, it is able to perform the uninterrupted UNRAVEL task in its entirety.

Figure 12 (left panel) shows the effect of varying RSI on the response time of Model A, with all parameters beyond those mentioned above set to the values given in Gilbert and Shallice (2002). The rates of each type of sequence error as a function or RSI are shown in figure 12 (right panel). It is clear from figure 12 that while Model A successfully reproduces the observed effect of RSI on response time (cf. figure 12, left panel, and figures 2, 5, and 8, right panels), it also produces an effect of RSI on sequence errors in the non-interrupted UNRAVEL task, with more perseverative errors at short RSI than at long RSI. The latter is inconsistent with our empirical results.

The effect of RSI on sequence errors in the non-interrupted UNRAVEL task is a result of carryover of task demand activation from one trial to the next – a feature of the Gilbert and Shallice (2002) model that the authors include in order to account for the effects of differential task difficulty and item-specificity on switch costs. In their model, a small proportion (20%) of task demand unit activity is retained from one trial to next. In the context of the UNRAVEL task this means, for example, that at the beginning of, say, a 'V' trials the activity of the 'A' task demand unit will be initialized to approximately 0.2 (rather than 0.0), given that the 'A' task demand unit will have been near its maximum (of 1.0) at the end of the previous trial. This predisposes Model A to perseverative errors. Note in particular that this activity is more disruptive when the RSI

[10] The negative bias on units effectively implements tonic inhibition. More negative bias is required to prevent over-excitation in the system when additional task demand and output units are added.

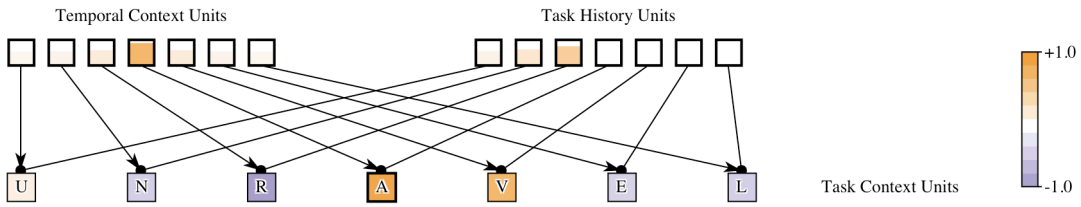


Figure 13: Place-keeping components of Model B, where the Task Context Units are identified with the corresponding Top Down Control Input Units of Model A (see figure 11). Temporal Context units excite corresponding Task Context units, while Task History units inhibit corresponding Task Context units. The strength of connections is shown in table 4. The shading of the units shows their activity when the system is about to perform the 'A' subtask (immediately after having successfully completed the 'R' subtask).

is small than when it is large. Note also that the model produces negligible rates of anticipatory errors at all values of RSI. Increasing task demand carryover yields a gradient of perseverative errors, with P2 errors also occurring but at lower rates than P1 errors, while decreasing it reduces the number of perseverative errors at all values of RSI. Setting task demand carryover to 0% (instead of 20%) completely eliminates sequence errors of all types in Model A's behavior.

6.2 Augmenting the Model with a Place-Keeping Mechanism

To model place-keeping we introduce two additional banks of units, as shown in figure 13 – a bank of Temporal Context units, which maintain a noisy representation of the current place in a sequential task, and a bank of Task History units, which maintain a memory of recently completed tasks. These combine to activate Task Context units, which are equated with the Top Down Control Input units of figure 11. We calculate activation of units in the three banks as follows:

In the case of Temporal Context units, which have one-to-one connections with their corresponding Task Context units, activation at the beginning of a block of UNRAVEL trials is initialized to:

$$TC_i = \begin{cases} \alpha, & \text{if } i = 1 \text{ (i.e., the first task)} \\ \alpha \cdot (1-\alpha)^{\sqrt{d}}, & \text{where } d \text{ is the number of steps between } i \text{ and the first task} \end{cases}$$

In this equation, α is a parameter (which we refer to as the *temporal acuity*) which controls the "sharpness" of the Temporal Context signal. When α is 1.0, the first Temporal Context unit is fully activated and all others have activation of 0.0. When α is less than 1.0, activation is concentrated on the first unit, but nearby units are also slightly active. For example, when α is 0.75, unit 1 will be set to 0.750, units 2 and 7 will be set to 0.106, units 3 and 6 will be set to 0.068, and units 4 and 5 will be set to 0.047.[11] After each step (i.e., immediately following the generation of a response and in

[11] The square root factor in the equation is designed to ensure activity of nearby units does not drop off too quickly with increasing distance.

Parameter	*Value*
Temporal acuity (α)	0.75
Task History persistence (ρ)	0.50
Temporal Context to Task Context weight	50.0
Task History to Task Context weight	15.0
Task Context noise (s.d.)	0.050
Task Context lateral inhibition	0
Task Context selection threshold	0.90
Step size	0.015
Task update probability (υ)	0.99

Table 4: Place-keeping parameters and their values used to generate figure 14.

preparation for the next step), the Temporal Context signal is shifted one unit to the right.

In the case of Task History units, activations are initialized to zero, but following each step of the task (i.e., immediately following the generation of a response and in preparation for the next step), the unit corresponding to that task is set to a value given by the parameter ρ (*Task History persistence*), and the activation of all other Task History units is multiplied by ρ. Thus, if ρ is 0.50, after completion of the first step the first Task History unit will be set to 0.50, while all others will be zero. After the second step, the first Task History unit will be 0.25, the second will be 0.50, and all others will be zero, and so on.

In the case of Task Context units, activation is calculated according to the interactive activation equations used in Model A (i.e., that of Gilbert and Shallice, 2002; see the appendix). As in that model, activation calculation is governed by a set of parameters. A complete listing of the parameters and their values for the place-keeping component of the model are given in table 4. Beyond those already discussed, the place-keeping component includes a noise parameter (whose value is larger than that in the set-shifting network), a step size parameter (whose value is the same as that in the set-shifting network), and a response threshold. In contrast to the set-shifting mechanisms, we assume selection of a Task Context unit (and subsequent excitation of the appropriate Task Demand unit in Model A) occurs when a Task Context unit's activation exceeds a threshold of 0.90. This typically yields a selection decision in 40 to 60 processing cycles.

One final parameter is the task update probability (υ). This is the probability that, following completion of a subtask, updating of Temporal Context and Task History units takes place. In other words, we assume that participants occasionally fail to update their place-keeping representations following generation of a response and in preparation of the next UNRAVEL step.

The graphs in figure 14 show the effect of the proposed place-keeping mechanisms on mean baseline response time (left panel) and the different types of sequence error (right panel) for three levels of response-stimulus interval. These graphs were generated with

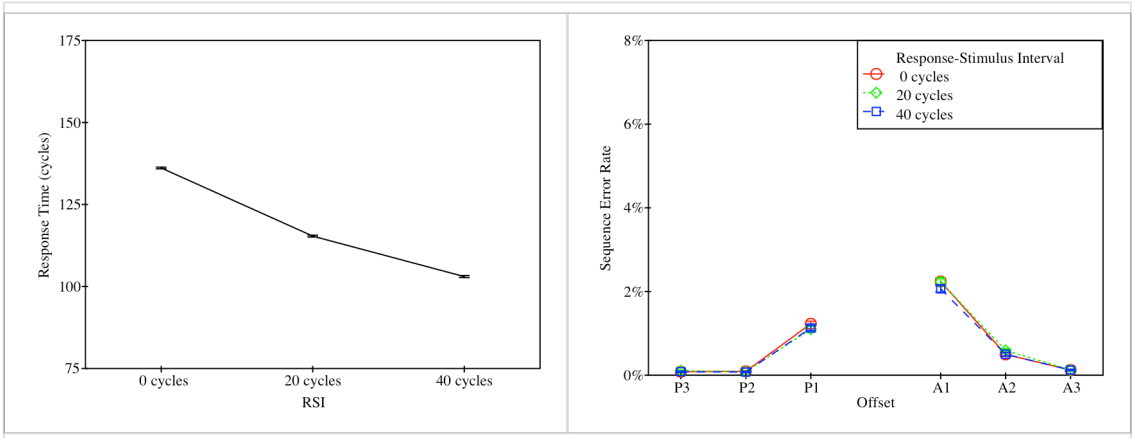


Figure 14: Predictions of Model B. Left panel: Effect of RSI on response time (measured in processing cycles). Right panel: Effect of RSI on the percentage of different types of sequence error in uninterrupted trials of the UNRAVEL task. The data are based on 40 virtual participants, with parameters set as in the simulation with Model A (see figure 12), except task-demand carryover was decreased to 10% Error-bars represent one standard error about the mean.

the place-keeping parameters set to the values given in table 4 and all aspects of Model A as in the earlier simulations. Note in particular that the full model (without considering interruption trials) produces the requisite decrease in response times as response-stimulus interval increases, together with gradients over both anticipatory and perseverative errors that are independent of response-stimulus interval. Moreover, while the Temporal Context units provide excitation to Task Context units that is symmetric about the current task, the inhibitory connections from Task History units serve to bias the system slightly towards anticipatory, and away from perseverative, sequence errors.

6.3 Modeling the Effect of Interruptions

We have proposed that the effect of an interruption is to degrade the representations of Temporal Context and Task History. Within the model there are multiple ways in which each of these might be achieved. Figure 15 shows the effect (for different values of response-stimulus interval) on sequence errors of decreasing temporal acuity (α) in the Temporal Context units and decreasing the persistence (ρ) of Task History units. In broad terms the left panel shows the simulated effect of short interruptions while the right panel shows the simulated effect of long interruptions (cf. figure 14, right panel). Decreasing the acuity of the Temporal Context signal results in less focused activation of the correct Task Context unit, and an increase in activation of nearby Task Context units. Decreasing the persistence of Task History units lessens their suppressing effect on perseverative errors. When that effect is sufficiently reduced (figure 15, right panel), perseverative errors dominate anticipatory errors, largely because of the model's occasional failure to advance steps, as given by the υ parameter. Indeed, similar effects may be obtained by keeping persistence fixed and decreasing υ. The model therefore does not directly speak to the issue of whether interruptions affect the representation of subtask history or the updating of position in task.

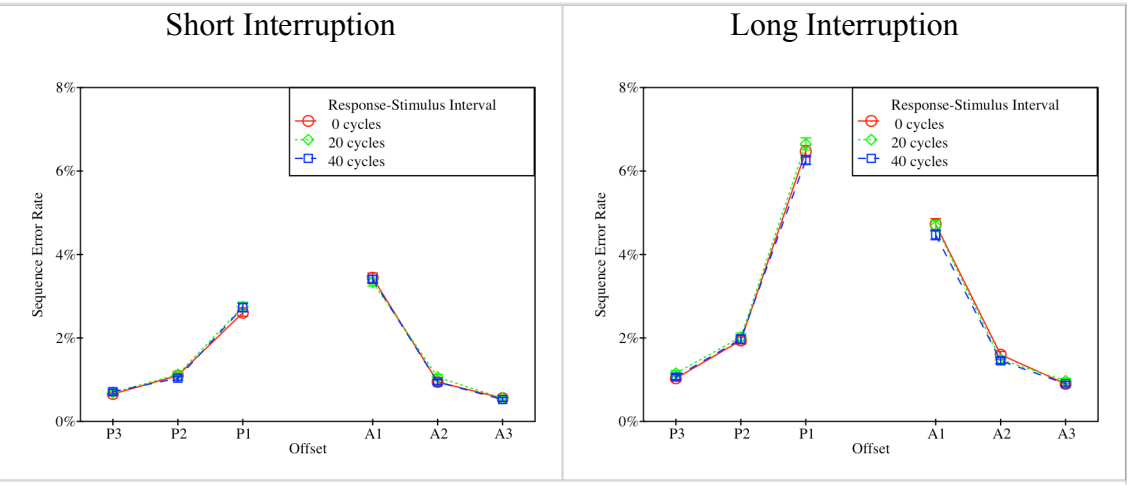


Figure 15: Predictions of Model B (i.e., Model A plus the place-keeping mechanism). Left panel: Simulation of short interruptions for three different response-stimulus intervals (α = 0.70, ρ = 0.25). Right panel: Simulation of long interruptions for three different response-stimulus intervals (α = 0.65, ρ = 0.00). Note that in neither case is there any appreciable effect of response-stimulus interval. Error-bars represent one standard error about the mean.

7 General Discussion

We have presented a model of the UNRAVEL task that combines the proposed set-shifting model of Gilbert and Shallice (2002) with a proposed place-keeping mechanism that embodies five assumptions:

1. Temporal context is maintained through a noisy or somewhat fuzzy representation of one's place in a task. During uninterrupted completion of a task this representation is updated as the task progresses in response to subtask completion.
2. A separate representation of subtask history (i.e., recently completed subtasks) works to suppress perseverative errors. This representation is also event-based, in the sense that subtask completion results in updating of the record of subtask history. It is also subject to decay.
3. Selection of the appropriate subtask at any point is a function of both of the above representations, and it is this selection that results in the all-or-none setting of top-down activation to lower-level mechanisms that control task set.
4. Updating of temporal context and subtask history units on completion of a subtask is fallible.
5. Interruptions result in blurring of the temporal context signal and decreased effectiveness of the subtask history record, with longer interruptions having larger effects on each representation.

These assumptions work together to produce the observed behavior. The assumptions provide an alternative to the Altmann and Trafton (2015) model of place-keeping in that rather than attributing perseverative and anticipatory errors to failures in separate mechanisms (the remember and advance mechanisms, respectively), such errors are attributed instead to a noisy representation of position, with subtask history working to suppress perseverative errors.

The above points specifically concern the mechanism (or mechanisms) that support place-keeping, yet a key empirical finding is that place-keeping in the UNRAVEL task is not affected by the response-stimulus interval. In fact, as shown in figure 12 (right panel), response-stimulus interval does affect error type in Model A. This is due entirely to the carryover of activation in the model's Task Demand units, as decreasing the carryover of activation from 20% to 0% (or to even just 10%) eliminates the effect. In fact, the inclusion of task carryover in models of set-shifting is contentious. Gilbert and Shallice (2002) argue for its inclusion in order to explain effects of differential task difficulty on task switch costs (with switch costs being, paradoxically, higher for the transition from hard to easy tasks, than from easy to hard tasks). While carryover of task set activation can account for the effect, Yeung and Monsell (2003) did not find paradoxical switch costs in a study where the subtasks used different response sets (as is the case in the UNRAVEL task). Moreover the level of carryover employed in the simulations of Gilbert and Shallice is somewhat arbitrary, with their appendix suggesting that the results hold for any level between 0% (i.e., no task demand carryover!) and 37%.

Our specific mathematical formulation of place-keeping, in terms of an activation-based mechanism conforming to the same dynamics as the Gilbert and Shallice (2002) set-shifting model, also warrants some discussion. The mathematical formulation merely provides a way of combining the two key sources of information – temporal context and subtask history – in the selection of the current subtask. We hold no strong commitment to the specifics of the underlying formulae. Indeed, some aspects of the mathematical formulation are problematic (such as the use of negative activations and their interaction with lateral inhibition), and the combination of information sources might equally be expressed purely in probabilistic terms, yielding a profile of the probability of selection of each subtask during the interrupted UNRAVEL task. In our other words, our mathematical formulation is intended to be illustrative or demonstrative, rather than definitive.

Moreover, while we have adopted the Gilbert and Shallice (2002) model as the set-shifting substrate, this is also for expository purposes. Other models of set-shifting have been proposed (e.g., Brown et al., 2007; Grange et al., 2013; Sexton & Cooper, 2017, as well, of course, as Altmann & Gray, 2008). Our choice to use the Gilbert and Shallice model was motivated primary by the simplicity of its proposed mechanisms. This simplicity ensures that the operation of our proposed place-keeping mechanisms is not obscured.

Finally, we have also described the place-keeping mechanism as a single system. That is, we have described Temporal Context units and Task History units as if they form an integrated mechanism. While our proposal is that temporal context and task history interact in the selection of position in a multi-step task, the latter may well be retained in a general purpose working memory, on the assumption that such a memory is subject to generally accepted memory biases of recency and decay.

8 Conclusion

The UNRAVEL task of Altmann and colleagues has previously proven to be well-suited to the empirical investigation of interruptions. We have used the task in a somewhat different way, focusing on the potential interaction between two cognitive control processes – set-shifting and place-keeping – in the performance of a relatively

complex task. To summarize, we found a decrease in response time on the UNRAVEL task with increasing response-stimulus interval. We interpret this as a fingerprint of set-shifting. Equally, we found increased rates of perseverative errors with longer interruptions on the UNRAVEL task. We similarly interpret this as a fingerprint of Altmann et al.'s (2014) place-keeping function. However throughout three experiments we found no evidence to suggest an interaction between these functions. Moreover, the behavioral patterns were similar across three experiments using different interrupting tasks, strongly implying that the effects are independent of the interrupting task and hence that they specifically relate to cognitive control.

To account for our results we have presented a model in which the processes responsible for set-shifting and place-keeping operate at separate levels and on distinct task representations. Effects related to set-shifting arise at the sub-task level, where an extended response-stimulus interval allows preparation, presumably in the form of task-set reconfiguration. Effects related to place-keeping arise at the superordinate level, where a noisy sense of position in the task, combined with a potentially unreliable mechanism for updating that sense and an imperfect memory of completed subtasks, may give rise to perseverative or anticipatory errors.

Acknowledgements

We are grateful to Erik Altmann, Nick Sexton, André Vandierendonck and several anonymous reviewers for comments on earlier drafts of this manuscript.

Appendix: Additional Details of the Models

Full code for both models, written in the C programming language, is available from http://www.ccnl.bbk.ac.uk/models. This appendix provides a verbal description of that code.

Model A and Model B both function according to standard interactive activation principles. Thus, in both models all units have an associated activation value that varies over time between a minimum value (*min* = –1) and a maximum value (*max* = +1). The models operate on each trial by initializing the activation of all units for time $t = 0$, as described below, and then repeatedly updating activations of all units at successive time increments ($t = 1$; $t = 2$; $t = 3$; …), again as described below, until an output unit's activation meets a response criterion. The specific response is that which corresponds to the output unit that meets the response criterion, while the number of time cycles between stimulus presentation (which need not occur at $t = 0$) and reaching the response criterion is an index of the model's response time.

Activation Initialization

At the beginning of each trial, the activation of each unit (in both models) is set as follows:

For each of the 14 input units:

$$act_{in,i}(0) = 0$$

For each of the 14 output units:

$$act_{out,i}(0) = \begin{cases} step \times \beta_{out} \times max & if\ \beta_i > 0 \\ -step \times \beta_{out} \times min & if\ \beta_i < 0 \end{cases}$$

where *step* determines the granularity of time steps and is set to 0.0015 as in Gilbert and Shallice (2002) and β_{out} is the bias on output units (set to –9 for all such units). For these parameter values this results in all output units being initialized to –0.0135. This is the value that output units would approach, given the step size and bias, in the absence of any input.

For each of the 7 task demand units:

$$act_{td,i}(0) = carryover \times prev_i$$

where *carryover* is a parameter that determines the strength of task demand carry over from one trial to the next (set at 0.2, as in Gilbert & Shallice, 2002) and $prev_i$ is the activation of task demand unit i at the end of the previous trial (or zero, for the first trial in a block).

For each of the 7 top-down control units of Model A and each of the 7 task context units of model B:

$$act_{tc,i}(0) = 0$$

In addition, for Model B:

Each temporal context unit is initialized at the beginning of a block as described in the main text, i.e.:

$$act_{tmp,i}(0) = \begin{cases} \alpha, & \text{if } i = 1 \text{ (i.e., the left-most unit)} \\ \alpha \cdot (1-\alpha)^{\sqrt{d}}, & \text{where } d \text{ is the number of steps between } i \text{ and unit 1} \end{cases}$$

and each of the 7 task history units is initialized to zero at the beginning of each block:

$$act_{th,i}(0) = 0$$

The activation of temporal context and task history units is fixed throughout each trial but immediately after generation of a response to task *i* and effectively at the beginning of the subsequent trial, the activation of temporal context units is shifted by one unit to the right (cycling back to the left-most unit as if the units were arranged in a circle rather than linearly), the activation of task history unit *i* is set to ρ, and the activation of each other task history unit is multiplied by ρ, thus implementing task history decay.

Activation Calculation

Following Gilbert and Shallice (2002), each non-input unit's activation value is updated on each processing cycle according to the following equations:

$$act_{x,i}(t+1) = act_{x,i}(t) + \varepsilon + step \times net_{x,i}(t) \times \begin{cases} max - act_{x,i}(t) & if\ net_{x,i}(t) > 0 \\ act_{x,i}(t) - min & if\ net_{x,i}(t) < 0 \end{cases}$$

where $act_{x,i}(t)$ is the activation of unit *i* of type *x* at time *t*, ε is a random noise term drawn from a Gaussian distribution with mean of 0 and standard deviation of 0.006, *step* controls the step size as described above, $net_{x,i}(t)$ is the net input to unit *i* of type *x* at time *t* as described below, and *min* and *max* are parameters that determine the activation range (fixed at –1 and +1 respectively, as noted above). Following this calculation, all activations are clipped so that they do not fall outside of the range defined by *min* and *max*.

The activation of input units remains at zero throughout the preparation interval (which in the simulations presented here, is either 0, 20 or 40 cycles), after which the input units corresponding to the trial's stimulus dimensions (e.g., *italic*, *near*, *red*, *above*, *vowel*, *even*, *more*, for the left-most stimulus in figure 1 of the main text) are set to *max*, while all others remain at zero.

Activation Propagation on Each Processing Cycle

For each input unit activation is set to zero or *max*, as described in the previous subsection. For all other units whose activation varies throughout a trial, that variation is determined by $net_{x,i}(t)$, the net input to unit *i* of type *x* at time *t*. This is calculated separately for each type of unit and in addition may depend on whether a top-down control unit (for Model A) or task context unit (for Model B) is selected. $net_{x,i}(t)$ is calculated for each type of unit as follows:

For each of the 7 top-down control units (in Model A) or task context units (in Model B):

$$net_{tc,i}(t) = \beta_{tc} + \sum_j act_{tmp,j}(t-1) \cdot w_{tc,ji} - act_{th,i}(t-1) \cdot w_{hc,i}$$

where

- β_{tc} is the bias on top-down control / task context units, set to –4 in both cases,

- $act_{tmp,j}(t-1)$ is the activation of the j^{th} temporal context unit at time $t-1$ (or zero for Model A), $w_{tc,ji}$ is the weight from temporal context unit j to top-down control / task context unit i, and the sum represents the excitation from temporal context to the relevant control unit, and
- $act_{th,i}(t-1)$ is the activation of the i^{th} task history unit at time $t-1$ (or zero for Model A), $w_{hc,i}$ is the weight from task history unit i to top-down control / task context unit i, and the term represents the inhibition from task history on the relevant control unit.

For each of the 7 task-demand units (in both models):

$$net_{td,i}(t) = \beta_{td} + s_i - \lambda \sum_{j}^{j \neq i} act_{td,j}(t-1) + \sum_{k} s_{td,ki} \cdot act_{out,k}(t-1)$$

where

- β_{td} is the bias on task demand units, set to –10,
- s_i is the top-down control input strength for task demand unit i, if the top down control unit (for Model A) or task context unit (for Model B) for task i is selected, and zero otherwise,
- λ is the lateral inhibition parameter, set to 4.0, and the sum represents the lateral inhibition from other task demand units, and
- $s_{td,k}$ is the strength of connections from output unit k to task demand unit i (which is +1 for connections from the two output units for task i and –1 for all other output units), and the sum (over all 14 output units) represents the recurrent excitation / inhibition of output units on task demand units.

For each of the 14 output units (in both models), activations are held fixed during the preparation interval. Only upon stimulus presentation are output units updated, with the net input to such units calculated as:

$$net_{out,i}(t) = \beta_{out} - \lambda \cdot act_{out,i'}(t-1) + s_{io,i} \cdot act_{in,i}(t-1) + \sum_{j} s_{to,ji} \cdot act_{td,j}(t-1)$$

where

- β_{out} is the bias on output units, set to –9,
- λ is the lateral inhibition parameter (set to 4.0, as above), i' is the other output node for this task (e.g., if i corresponds to the "above" output node, then i' will be the "below" output node, and *vice versa*), and the second term in the equation represents lateral inhibition between competing output nodes,
- $s_{io,i}$ is the strength of task-specific input-output channels, and the third term in the equation represents excitation of output units from corresponding input units, and
- $s_{to,ji}$ is the strength of association between task-demand and output units (assumed to be 2.5 for output units associated with the specific task and –2.5 for output units associated with each other task), and the final term represents excitation /inhibition of output units from task demand units.

Task Selection

For Model A, the appropriate top down control unit for each UNRAVEL task was selected after 50 cycles, corresponding to a system with perfect place keeping but where recall of position in the UNRAVEL sequence takes 50 processing cycles. For Model B, a task control unit was selected if and when the activation of that unit exceeded a threshold, set to 0.9. If more than one unit's activity exceeded the threshold only the most active of those units was selected. With the parameters given here, selection in Model B typically took between 20 and 40 processing cycles.

The Response Criterion

On each cycle, and after the activations of all units have been updated, the difference between the activation of the most active and next most active output units is calculated. If that difference exceeds a threshold – set to 0.15 as in the model of Gilbert and Shallice (2002) – then processing for the trial is terminated and the response corresponding to the output unit with the highest activation is produced.

Interruptions

Interruptions of each duration and at each level RSI were modeled as follows. First, the model was allowed to cycle for a number of cycles corresponding to the RSI (reflecting the fact that interruptions in the experimental tasks occurred after the RSI interval). The activation of all task history units was then decreased by a decay factor that varied with interruption length (0.5 for short interruptions, 0.0 for long interruptions). The activation of task context units was reset to zero. The activation of all temporal context units was reset with decreased temporal acuity (of 0.70 for short interruptions and 0.65 for long interruptions). Finally, if a task context unit has been selected during the preparation interval, this was deselected. The UNRAVEL trial then resumed, with immediate presentation of an UNRAVEL stimulus, as in the experimental task.

References

Allport, A., & Wylie, G. (2000). Task-switching, stimulus–response bindings, and negative priming. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 36–70). Cambridge, MA: MIT Press.

Altmann, E. M., & Gray, W. D. (2008). An integrated model of cognitive control in task switching. *Psychological Review*, *115*(3), 602-639.

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin and Review*, *14*, 1079–1084.

Altmann, E. M., & Trafton, J. G. (2015). Brief Lags in Interrupted Sequential Performance: Evaluating a Model and Model Evaluation Method. *International Journal of Human-Computer Studies*, 79, 51-65.

Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General*, 143(1), 215.

Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: Evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review*, *12*(6), 1011-1017.

Brown, J. W., Reynolds, J. R., & Braver, T. S. (2007). A computational model of fractionated conflict-control mechanisms in task-switching. *Cognitive Psychology*, *55*(1), 37-85.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551-581.

Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science*, *21*, 31–61.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332-361.

Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology*, *17*, 297–338.

De Jong, R. (2000) An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance XVIII* (pp. 357–376). Cambridge, MA: MIT Press.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology*, *44*(3), 297-337.

Grange, J. A., Juvina, I., & Houghton, G. (2013). On costs and benefits of n− 2 repetitions in task switching: Towards a behavioural marker of cognitive inhibition. *Psychological Rresearch*, *77*(2), 211–222.

Hambrick, D. Z., & Altmann, E. M. (2015). The role of placekeeping ability in fluid intelligence. *Psychonomic Bulletin and Review*, 22(4), 1104-1110.

JASP Team (2018). JASP (Version 0.8.5.1) [Computer software]

Koch, I., Gade, M., Schuch, S., & Philipp, A. M. (2010). The role of inhibition in task switching: A review. *Psychonomic Bulletin and Review*, *17*(1), 1-14.

Lashley, K.S. (1951). The problem of serial order in behaviour. In L. A. Jeffress (Ed.), *Cerebral Mechanisms in Behaviour* (pp. 112–136). New York: John Wiley.

Li, S. Y. W., Blandford, A., Cairns, P., & Young, R. M. (2008). The effect of interruptions on postcompletion and other procedural errors: An account based on the activation-based goal memory model. *Journal of Experimental Psychology: Applied*, *14*, 314–328.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*(2), 393–434.

Logie, R. H. (2016). Retiring the central executive. *The Quarterly Journal of Experimental Psychology*, *69*(10), 2093-2109.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1423–1442.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, *9*, 100–110.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, *7*(3), 134-140.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York, NY: Plenum Press.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231.

Schneider, D. W., & Logan, G. D. (2007). Defining task-set reconfiguration: The case of reference point switching. *Psychonomic Bulletin & Review*, *14*(1), 118–125.

Sexton, N. J. & Cooper, R. P. (2017). Task inhibition, conflict, and the n-2 repetition cost: A combined computational and empirical approach. *Cognitive Psychology*, *94*, 1–25.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., ... & Izukawa, D. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*(4), 388–402.

Trafton, J. G., Altmann, E. M., & Ratwani, R. (2011). A memory for goals model of sequence errors. *Cognitive Systems Research*, *12*, 134–143.

Vandierendonck, A., Liefooghe, B., & Verbruggen, F. (2010). Task switching: Interplay of reconfiguration and interference control. *Psychological Bulletin*, *136*(4), 601–626.

Washburn, M. F. (1916). Movement and mental imagery. Boston: Houghton Mifflin.

Watson, J. B. (1920). Is thinking merely the action of the language mechanisms? *British Journal of Psychology*, *11*, 86-104.

Yeung, N., & Monsell, S. (2003). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 455–469.